

ESCOLA SUPERIOR DE SAÚDE EGAS MONIZ

MESTRADO EM BIOLOGIA MOLECULAR EM SAÚDE



**ESTUDO DE REGIÕES DE
SUSCEPTIBILIDADE PARA O CANCRO DO
CÓLON E RECTO FAMILIAR DO TIPO X:
ANÁLISE DE GENES CANDIDATOS E DE
GANHOS/DELEÇÕES EM TUMORES**

**Tese de dissertação para obtenção do grau de Mestre em Biologia Molecular em
Saúde**

Gonçalo Filipe de Almeida Pereira

Setembro de 2014

ESCOLA SUPERIOR DE SAÚDE EGAS MONIZ

MESTRADO EM BIOLOGIA MOLECULAR EM SAÚDE



ESTUDO DE REGIÕES DE SUSCEPTIBILIDADE PARA O CANCRO DO CÓLON E RECTO FAMILIAR DO TIPO X: ANÁLISE DE GENES CANDIDATOS E DE GANHOS/DELEÇÕES EM TUMORES

**Tese de dissertação para obtenção do grau de Mestre em Biologia Molecular em
Saúde**

Gonçalo Filipe de Almeida Pereira
Orientado por Dra. Cristina Albuquerque

Setembro de 2014

Agradecimentos

A realização desta dissertação contou com importantes apoios sem os quais não se teria tornado uma realidade.

À Dra. Cristina Albuquerque, um agradecimento em especial pela disponibilidade, pelo saber transmitido, pelo esclarecimento de dúvidas ao longo deste trabalho e sem menos importância pela simpatia e boa disposição.

À Mestre Inês Francisco pela ajuda na revisão da tese e por todo o apoio no trabalho prático realizado.

À Professora Doutora Alexandra Maia e Silva pela preocupação e disponibilidade no esclarecimento de dúvidas.

À Dra. Patrícia Silva e ao Dr. Bruno Filipe agradeço a simpatia e a prontidão para ajudar sempre foi necessário.

À Dra. Sofia Fragoso e ao Mestre Hélio Belo, pela grande ajuda na realização do PCR quantitativo, quer no trabalho prático quer na análise dos resultados.

À Dra. Patrícia Machado e à Dra. Sidónia pela boa disposição e simpatia que demonstraram desde o primeiro dia.

À Doutora Branca Cavaco, Coordenadora da unidade de Investigação em Patobiologia Molecular, do Instituto Português de Oncologia, Lisboa, Francisco Gentil, E.P.E, por ter concedido a oportunidade de participar neste projecto.

Ao Serviço de Gastrenterologia, Clínica de Risco familiar e ao Serviço de Anatomia Patológica do Instituto Português de Oncologia, Lisboa, Francisco Gentil, E.P.E, pela disponibilização do material biológico e informação clínica dos doentes.

Às minhas colegas, Ana, Marisa e Filipa pela companhia e apoio ao longo do estágio. Dirijo um agradecimento especial à Filipa Fernandez pela integração no laboratório e pela prontidão em ajudar. Agradeço ainda a companhia em todos os lanches e almoços que partilhámos, sabendo que sem ela este estágio não teria sido a mesma coisa.

Aos restantes membros do Centro de Investigação em Patobiologia Molecular pelo ambiente acolhedor com que sempre me receberam.

E finalmente, aos meus pais, irmã e namorada pelo apoio, incentivo e compreensão demonstrado ao longo deste ano que me fizeram facilmente ultrapassar os tempos difíceis.

Resumo

O cancro do cólon e recto familiar do tipo X (FCCTX) define as famílias que preenchem os critérios de Amesterdão, mas nas quais não é identificada mutação germinal nos genes de reparação de erros de DNA do tipo *mismatch* (MMR) e cujos tumores não apresentam instabilidade de microsatélites. A sua causa molecular não é ainda conhecida. O presente trabalho teve como objectivo avaliar o envolvimento de genes localizados em duas regiões de susceptibilidade previamente identificadas (13q32-33 e 21q11) e o esclarecimento das perdas de heterozigotia (LOH) frequentes em 13q32-33 em tumores FCCTX, através de uma análise do número de cópias (*copy number*). Pretendeu-se ainda estudar o gene *BMPRIA* como gene de susceptibilidade para o FCCTX.

Foi efectuada análise de mutações germinais dos genes *HSPA13*, *SAMSN1* e *BMPRIA* em 15 indivíduos índice das famílias FCCTX e 2 familiares de uma das famílias. Foi ainda realizada, a nível germinal, a pesquisa de transcritos aberrantes dos genes *TEX30* e *SAMSN1* e quantificada a expressão destes e do *HSPA13* por qPCR em 3 indivíduos índice e 6 familiares de uma das famílias FCCTX. Da mesma forma procedeu-se também ao estudo da expressão diferencial dos transcritos alternativos do gene *TPP2*. Foi ainda avaliada a eventual patogenicidade de três mutações previamente detectadas no gene *SLC10A2*.

Não foram detectadas quaisquer mutações nem alterações de expressão potencialmente patogénicas nos genes estudados. As análises *in silico* e de segregação com a doença sugerem no entanto, uma mutação no gene *BMPRIA* e duas mutações no gene *SLC10A2*, para as quais são necessários estudos adicionais para concluir sobre a sua patogenicidade. Um dos transcritos do gene *TPP2* apresenta expressão diferencial entre amostras, o que sugere a continuação do seu estudo. A LOH frequente em 13q corresponde a ganho/amplificação.

Em conclusão, o presente estudo exclui os genes localizados em 21q11, *HSPA13* e *SAMSN1*, e em 13q32-33, *TEX30*, como possíveis genes candidatos para o FCCTX. Não é ainda possível excluir o gene *SLC10A2* nem o gene *TPP2*. Mutações no *BMPRIA* deverão ser raras em FCCTX. Parece existir um elevado grau de instabilidade genómica em lesões precoces FCCTX com ganhos/amplificações frequentes.

Palavras-chave: Cancro do cólon e recto familiar do tipo X, *HSPA13*, *SAMSN1*, *TPP2*

Abstract

Familial colorectal cancer type X (FCCTX) defines families that fulfill the Amsterdam criteria, but in whom no germline mutation is identified in the DNA mismatch repair (MMR) genes and whose tumors do not present microsatellite instability. Its molecular basis is not known yet. This study aimed to evaluate the involvement of genes located in two susceptibility regions previously identified (13q32-33 and 21q11) and clarification of the loss of heterozygosity (LOH) events in 13q32-33 in FCCTX tumors, by copy number analysis. *BMPR1A* gene was also studied as a susceptibility gene for FCCTX.

Analysis of germline mutations in *HSPA13*, *SAMSNI* and *BMPR1A* genes was carried out in 15 index individuals from FCCTX families and two relatives from one of the families. It was also performed, at the germline level, the analysis of aberrant transcripts of *TEX30* and *SAMSNI* genes and the expression of these genes and of *HSPA13* was quantified by qPCR in 3 index individuals and six relatives of one of the FCCTX families. Likewise the study of the differential expression of alternative transcripts of the *TPP2* gene was also performed. The eventual pathogenicity of three mutations previously detected in the *SLC10A2* gene was further evaluated.

We did not detect any potentially pathogenic mutations or changes in the expression of the genes studied. *In silico* analysis and segregation with disease suggest, however, a mutation in the *BMPR1A* gene and two mutations in the *SLC10A2* gene for which additional studies are required to complete pathogenicity evaluation. One of the *TPP2* transcripts presented differential gene expression between samples, suggesting the continuation of its study. Frequent LOH on 13q corresponded to gain/amplification.

In conclusion, the present study excludes genes located in 21q11, *HSPA13* and *SAMSNI*, and 13q32-33, *TEX30*, as candidate genes for FCCTX. The *SLC10A2* and *TPP2* gene cannot be excluded yet. Mutations in *BMPR1A* should be rare in FCCTX. Furthermore, there seems to be a high degree of genomic instability in early FCCTX lesions with frequent gains/amplifications.

Key-words: Familial colorectal cancer type X, *HSPA13*, *SAMSNI*, *TPP2*

Índice Geral

Agradecimentos	2
Resumo	4
Abstract	5
Índice Figuras	9
Índice Tabelas	12
Lista de Siglas	13
1. Introdução	15
1.1. Cancro do cólon e recto.....	15
1.2. Sequência adenoma-carcinoma	15
1.3. Síndromes hereditários de cancro do cólon e recto.....	17
1.3.1. Síndromes hereditários associados a polipose	17
1.3.2. Cancro do cólon e recto hereditário não associado a polipose	18
1.3.2.1. Síndrome de Lynch.....	19
1.4. Cancro do cólon e recto familiar do tipo X.....	21
1.4.1. Mapeamento de <i>loci</i> de susceptibilidade para FCCTX	22
1.4.2. Genes candidatos de susceptibilidade para o FCCTX na região 13q32-33	23
1.4.3. Genes candidatos de susceptibilidade para o FCCTX na região 21q11	25
1.4.3.1. <i>Microsomal Stress-70 Protein ATPase Core (HSPA13)</i>	26
1.4.3.2. <i>SAM Domain, SH3 Domain And Nuclear Localization Signals 1 (SAMSNI)</i>	28
1.4.4. <i>Bone morphogenetic protein receptor, type IA (BMPRIA)</i>	30
2. Objectivos	33
3. Materiais e métodos	34
3.1. Materiais.....	34
3.2. Métodos.....	36
3.2.1. Extracção de ácidos nucleicos	36
3.2.1.1. Extracção de DNA a partir de sangue periférico.....	36
3.2.1.2. Extracção de DNA a partir de tecido tumoral e mucosa normal incluído em parafina	36

3.2.1.3.	Extracção de RNA a partir de sangue periférico	36
3.2.2.	Amplificação do DNA por reacção em cadeia da polimerase (PCR)..	36
3.2.2.1.	Optimização das condições de PCR	37
3.2.2.2.	Análise electroforética do produto amplificado por PCR em gel de agarose	38
3.2.3.	Análise mutacional dos genes <i>HSPA13</i> , <i>SAMSN1</i> e <i>BMPRIA</i>	39
3.2.3.1.	Amplificação por PCR dos exões dos genes <i>HSPA13</i> , <i>SAMSN1</i> e <i>BMPRIA</i>	39
3.2.3.2.	Controlo da eficiência do produto amplificado por PCR em gel de agarose a 2% (p/v)	40
3.2.3.3.	Purificação dos produtos de amplificação por PCR	40
3.2.3.3.1.	Purificação enzimática dos produtos de amplificação por PCR	41
3.2.3.3.2.	Purificação em coluna dos produtos de amplificação por PCR utilizando o kit QIAquick® Gel Extraction (<i>Qiagen</i>).....	41
3.2.3.4.	Sequenciação automática.....	42
3.2.3.4.1.	Reacção de sequenciação.....	42
3.2.3.4.2.	Purificação e precipitação do DNA após a reacção de sequenciação	43
3.2.3.4.3.	Electroforese capilar	44
3.2.3.4.3.1.	Preparação das amostras para electroforese capilar	44
3.2.3.4.3.2.	Análise de resultados.....	44
3.2.4.	Análise de grandes deleções/inserções nos genes <i>TPP2</i> , <i>TEX30</i> e <i>SAMSN1</i>	44
3.2.4.1.	Reacção de transcrição reversa seguida de PCR (RT-PCR) para síntese de cDNA	45
3.2.4.2.	Amplificação por PCR do cDNA dos genes <i>TPP2</i> , <i>TEX30</i> e <i>SAMSN1</i>	46
3.2.4.3.	Análise dos transcritos alternativos identificados no contexto do FCCTX	48
3.2.4.4.	Análise da expressão do transcrito alternativo 13a identificado no gene <i>TPP2</i> através de PCR semiquantitativo	48
3.2.5.	<i>Real-Time</i> PCR (qPCR)	49
3.2.5.1.	Análise da expressão dos genes <i>TPP2</i> , <i>TEX30</i> , <i>SAMSN1</i> e <i>HSPA13</i> por <i>Real-Time</i> PCR	51
3.2.5.1.1.	Validação da eficiência de amplificação dos genes em estudo .	51
3.2.5.1.2.	Quantificação relativa dos genes em estudo	52

3.2.5.2.	Análise de <i>copy number</i> nos genes <i>TPP2</i> , <i>SAMSN1</i> e <i>HSPA13</i> nas famílias FCCTX por qPCR	53
3.2.5.3.	Análise de <i>copy number</i> na região 13q32-33 em tumores de indivíduos FCCTX por qPCR	54
3.2.6.	Análise das mutações c.292 G>A, c.868 C>T e c.511 G>T no gene <i>SLC10A2</i> no contexto do FCCTX	55
4.	Resultados	56
4.1.	Envolvimento de genes candidatos na susceptibilidade para o FCCTX.....	56
4.1.1.	<i>BMPRI1A</i>	56
4.1.2.	<i>HSPA13</i>	59
4.1.3.	<i>SAMSN1</i>	62
4.2.	Análise do envolvimento de genes candidatos localizados na região de susceptibilidade 13q32-33.....	67
4.2.1.	<i>TPP2</i>	67
4.2.2.	<i>TEX30</i>	75
4.2.3.	<i>SLC10A2</i>	77
4.3.	Análise de <i>copy number</i> na região 13q32-33 em tumores de indivíduos FCCTX.....	79
5.	Discussão	85
6.	Conclusão	93
7.	Referências bibliográficas	95
Anexos	101

Índice Figuras

Figura 1. Representação esquemática do modelo adenoma-carcinoma (Vogelgrama) com representação das alterações histológicas e moleculares mais frequentes.....	16
Figura 2. Representação esquemática dos dois transcritos codificantes do gene <i>HSPA13</i>	26
Figura 3. Representação esquemática dos domínios conhecidos da proteína <i>HSPA13</i> .	27
Figura 4. Representação esquemática da via de sinalização pró-apoptótica TRAIL, com representação das suas principais proteínas	27
Figura 5. Representação esquemática dos transcritos codificantes do gene <i>SAMSNI</i> ..	29
Figura 6. Representação dos domínios conhecidos da proteína <i>SAMSNI</i>	30
Figura 7. Representação esquemática da via de sinalização BMP com o ligando BMP e as principais proteínas da via.....	31
Figura 8. Representação esquemática do gene <i>BMPRIA</i> , <i>HSPA13</i> e <i>SAMSNI</i> , e respectiva localização dos <i>primers</i> utilizados.....	40
Figura 9. Representação esquemática do cDNA dos genes <i>SAMSNI</i> , <i>TEX30</i> , <i>TPP2</i> ...	47
Figura 10. Representação esquemática do método de fit points.....	49
Figura 11. Representação esquemática do método da segunda derivada.	50
Figura 12. Resultados da expressão relativa do transcrito wt do gene <i>HSPA13</i> para os indivíduos de 2 famílias FCCTX (L55 e L7).	60
Figura 13. Resultados da expressão relativa do transcrito wt do gene <i>HSPA13</i> em indivíduos saudáveis e em indivíduos com SL.....	61
Figura 14. Resultados da análise de copy number do gene <i>HSPA13</i> para os indivíduos de 2 famílias FCCTX (L55 e L7).....	62
Figura 15. Electroforese em gel de agarose 1,5% (p/v) dos produtos amplificados por PCR para o gene <i>SAMSNI</i> e 2% (p/v) para o gene <i>GAPDH</i>	64

Figura 16. Resultados da expressão relativa do transcrito wt do gene <i>SAMSN1</i> para os indivíduos de 2 famílias FCCTX (L55 e L7)	65
Figura 17. Resultados da expressão relativa do transcrito wt do gene <i>SAMSN1</i> em indivíduos saudáveis e em indivíduos com SL	66
Figura 18. Resultados da análise de copy number do gene <i>SAMSN1</i> para os indivíduos de 2 famílias FCCTX (L55 e L7).....	67
Figura 19. Electroforese em gel de agarose 1,2% (p/v) dos produtos amplificados por PCR para o transcrito alternativo 24a, transcrito alternativo 13a do gene <i>TPP2</i> e 2% (p/v) para o gene housekeeping <i>GAPDH</i> para os indivíduos de famílias FCCTX, indivíduos com SL e indivíduos saudáveis	69
Figura 20. Electroforese em gel de agarose 1,2% (p/v) dos produtos amplificados por PCR para o transcrito alternativo 13a do gene <i>TPP2</i> e 2% (p/v) para o gene housekeeping <i>GAPDH</i> para os indivíduos de famílias FCCTX, indivíduos com SL e indivíduos saudáveis	71
Figura 21. Resultados da expressão relativa do transcrito wt e do transcrito 13a do gene <i>TPP2</i> para os indivíduos de 2 famílias FCCTX (L55 e L7).....	72
Figura 22. Resultados da expressão relativa do transcrito wt e do transcrito 13a do gene <i>TPP2</i> em indivíduos saudáveis	73
Figura 23. Resultados da expressão relativa do transcrito wt e do transcrito 13a do gene <i>TPP2</i> em indivíduos com SL	74
Figura 24. Resultados da análise de copy number do gene <i>TPP2</i> para a região do intrão 6 e 21 em indivíduos de 2 famílias FCCTX (L55 e L7).....	75
Figura 25. Electroforese em gel de agarose 1,2% (p/v) dos produtos amplificados por PCR para o gene <i>TEX30</i> e 2% (p/v) para o gene housekeeping <i>GAPDH</i>	76
Figura 26. Resultados da expressão relativa do transcrito wt do gene <i>TEX30</i> para os indivíduos de 2 famílias FCCTX (L55 e L7).....	77
Figura 27. Resultados da análise de LOH para os marcadores de microssatélites D13S795 a D13S158	82

Figura 28. Resultados de *copy number* dos genes *DOCK9* e *TPP2* regiões do intrão 6 e 21 para os tumores e mucosas normais dos indivíduos das famílias FCCTX 83

Figura 29. Resultados da análise de *copy number* da região 13q32-33 através da análise dos genes *DOCK9* e *TPP2*..... 84

Índice Tabelas

Tabela 1. Critérios de Amesterdão I e II.....	18
Tabela 2. Critérios de Bethesda	20
Tabela 3. Principais características das famílias de síndrome de Lynch e FCCTX	21
Tabela 4. Programa utilizado na reacção de sequenciação	43
Tabela 5. Caracterização das mutações identificadas no gene <i>BMPRIA</i>	58
Tabela 6. Caracterização das mutações identificadas no gene <i>HSPA13</i>	59
Tabela 7. Caracterização das mutações identificadas no gene <i>SAMSN1</i>	63
Tabela 8. Caracterização das mutações identificadas no gene <i>SLC10A2</i>	78

Lista de Siglas

ALK3 – Activin-like kinase 3

APAF – Polipose adenomatosa familiar atenuada

APC – Adenomatous poliposis coli

BMP – Bone morphogenetic protein

BMPRIA – Bone morphogenetic protein receptor, type IA

CA – Critérios de Amesterdão

CCR – Cancro do cólon e recto

CIN – Instabilidade cromossómica

Cp – Crossing points

DEPC – dietilpirocarbonato

DOCK9 – Dedicator of Cytokinesis 9

E – Eficiência

FADD – FAS-associated protein with death domain

FCCTX – Cancro do colon e recto familiar do tipo X

FFPE – Tecido incluído em parafina

FSP – Frameshift peptides

GAPDH – Glyceraldehyde 3-phosphate dehydrogenase

HACS1 – Hematopoietic Adapter-Containing SH3 and Sterile α -motif (SAM) domains 1

HDAC1 – Histona deacetilase 1

HNPPC – Cancro do colon e recto hereditario não associado a polipose

Hsp70 – Heat shock protein 70

HSPA13 – Microsomal Stress-70 Protein ATPase Core

LOH – Perda de heterozigotia “Loss of heterozygosity”

LYN – Proteína tirosina cinase Lck/Yes-related novel

METTL21C – Methyltransferase Like 21C

MGMT – Proteína metilguanina metiltransferase

MHC-I – Complexo principal de histocompatibilidade classe I

MMR – Mismatch repair

MSI – Instabilidade de microsatélite

MSI-H – Instabilidade de microsatélite -*High*

MSI-L – Instabilidade de microsátélite -Low
 N – Grupo controlo
NRIP1 – *Nuclear Receptor Interacting Protein 1*
 NS – Não significativo
 PAF – Polipose adenomatosa familiar
 PAM – Polipose adenomatosa associada ao gene *MUTYH*
 qPCR – PCR quantitativo / *Real time-PCR*
RPS13 – *proteína ribossomal S13*
SAM – *Sterile α -motif*
SAMSN1 – *SAM Domain, SH3 Domain And Nuclear Localization Signals 1*
SDHA – *sucinato desidrogenase subunidade A*
SH3 – *SRC Homology 3*
 SL – Síndrome de Lynch
SLC10A2 – *Solute Carrier Family 10 (Sodium/Bile Acid Cotransporter Family), Member 2*
 SNP – Polimorfismo de nucleótido único
STCH – *Stress 70 Protein Chaperone*
STK24 – *Serine/Threonine Kinase 24*
STS – *Sequence tagged-site*
TEX30 – *Testis Expressed 30*
TGF- β – *Transforming growth factor*
TPP2 – *Tripeptidyl Peptidase II*
TRAIL – *TNF-related Apoptosis-inducing ligand*
 wt – *Wild type*

1. Introdução

1.1. Cancro do cólon e recto

O cancro do cólon e recto (CCR), encontra-se descrito como o terceiro cancro mais diagnosticado em homens e o segundo em mulheres. A incidência acentua-se na Austrália e Nova Zelândia, bem como na Europa e América do Norte, sendo que a menor incidência se encontra em África, demonstrando assim a associação entre este tipo de cancro e os países desenvolvidos (Jemal et al. 2011).

Sendo uma doença multifactorial, os factores de risco associados a esta podem ter uma componente ambiental e/ou genética. Os principais factores de risco já estudados são, o tabagismo, o sedentarismo, a obesidade, o consumo excessivo de carnes vermelhas e o alcoolismo (Jemal et al. 2011). Cerca de 70% dos casos de CCR correspondem a formas esporádicas, enquanto, os restantes 30%, possuem uma história familiar da doença. Destes, cerca de 5-6% encontram-se associados a síndromes hereditárias de CCR (Vasen et al. 2008).

1.2. Sequência adenoma-carcinoma

Em 1990 foi proposto o primeiro modelo de tumorigénese colorectal, a sequência adenoma-carcinoma ou Vogelgrama (Figura 1), a qual descreve um conjunto de alterações histológicas acompanhadas de alterações moleculares bem caracterizadas. Estas alterações ocorrem em genes específicos, genes supressores de tumor e oncogenes, desde a iniciação à progressão tumoral (Pino & Chung 2010).

De acordo com este modelo, a alteração mais precoce corresponde à inactivação do gene *adenomatous polyposis coli* (*APC*), a qual conduz à formação de um adenoma a partir da mucosa cólica normal. A proteína APC é uma proteína multifuncional mas tem como principal função de supressão tumoral a regulação negativa da β -catenina na via de transdução de sinal WNT/ β -catenina (Saito-diaz et al., 2013).

A activação da via WNT/ β -catenina, (Figura 1), implica a translocação da β -catenina para o núcleo, onde esta actua como factor de transcrição, activando genes envolvidos na proliferação e diferenciação celular. Em condições normais o complexo APC, Axina e GSK3 β actua de forma a regular os níveis citoplasmáticos de β -catenina, levando à sua degradação no proteossoma. Quando ocorrem mutações nos genes *APC* ou *CTNNB1* que codifica para a β -catenina, a via permanece activada

constitutivamente levando a uma proliferação descontrolada da mucosa cólica o que resulta na formação de um adenoma precoce (Pandurangan 2013; Saito-diaz et al. 2013).

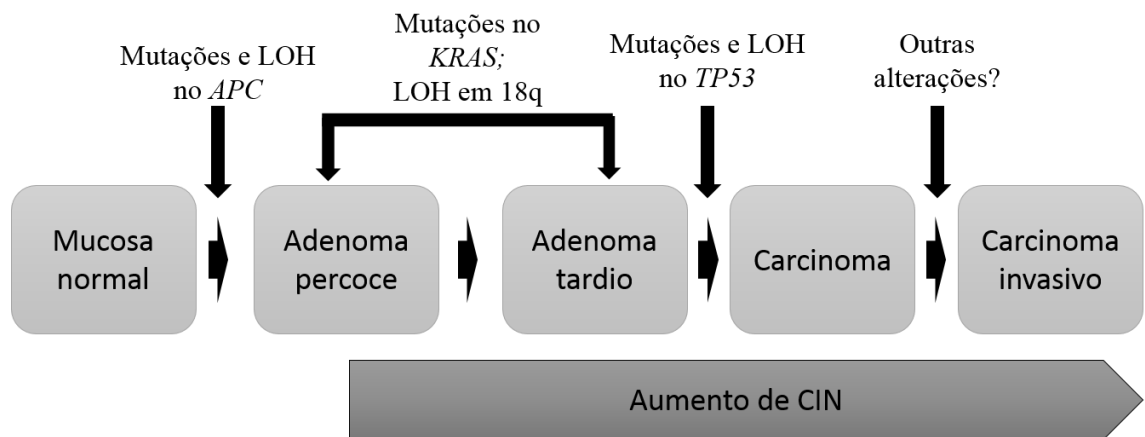


Figura 1. Representação esquemática do modelo adenoma-carcinoma (Vogelgrama) com representação das alterações histológicas e moleculares mais frequentes (adaptado de (Pino & Chung 2010)), LOH (*Loss of heterozygosity*); CIN (Via da instabilidade cromossômica).

Em seguida, ocorrem mutações no oncogene *KRAS* e perda de heterozigotia (LOH) em genes da via de sinalização TGF- β , nomeadamente no gene *SMAD2* e *SMAD4* (18q), contribuindo para a progressão tumoral. Posteriormente, a inativação do gene supressor de tumor *TP53* leva à transição de adenoma tardio para carcinoma (Pino & Chung 2010).

A estas alterações genéticas associam-se também alterações na segregação cromossômica, consequência entre outros factores também da inativação do gene *APC*, resultando em aneuploidias, ganhos e/ou perdas de porções cromossômicas, que se traduzem em perdas de heterozigotia, as quais são características da via da instabilidade cromossômica (CIN). Em 1993, foi proposta uma outra via de instabilidade genética, a instabilidade de microssatélites (MSI), a qual é seguida por cerca de 15% dos tumores esporádicos e pela quase totalidade dos tumores que ocorrem no contexto do síndrome hereditário de CCR mais frequente, o síndrome de Lynch (Pino & Chung 2010).

1.3. Síndromes hereditários de cancro do cólon e recto

Os síndromes hereditários de CCR, são definidos com base num diagnóstico clínico e/ou genético, e dividem-se em duas categorias principais, os síndromes hereditários associados a polipose e o cancro do cólon e recto hereditário não associado a polipose (HNPCC) (Jasperson et al. 2011).

1.3.1. Síndromes hereditários associados a polipose

Os síndromes hereditários associados a polipose englobam, a polipose adenomatosa familiar (PAF), a polipose adenomatosa familiar atenuada (APAF), a polipose adenomatosa associado ao gene *MUTYH* (PAM), a polipose serreada e a polipose hamartomatosa (Jasperson et al. 2011).

A PAF apresenta uma prevalência de 1 em 10.000 indivíduos, apresentando como características clínicas, uma transmissão autossómica dominante e o desenvolvimento de centenas a milhares de pólipos no cólon. A idade média de diagnóstico é de 39 anos e molecularmente é caracterizada por mutações germinais no gene *APC*. Mutações neste gene são também responsáveis pela APAF que apresenta um menor número de pólipos (≤ 100 pólipos), sendo que, em ambos os síndromes a localização da mutação germinal ao longo do gene está associada à agressividade do fenótipo no cólon bem como à presença variável de manifestações extra-cólicas (Jasperson et al. 2011).

Relativamente à polipose adenomatosa associada ao gene *MUTYH*, este é um síndrome de transmissão autossómica recessiva, definido com base genética por identificação de mutações germinais bialélicas no gene *MUTYH*. Fenotipicamente, este síndrome mimetiza a APAF com o desenvolvimento de 10-100 pólipos por todo o cólon, sendo o diagnóstico desta patologia efectuado numa idade mais tardia (Jasperson et al. 2011).

Para além destes, existem ainda dois síndromes menos frequentes, o síndrome de polipose hamartomatosa que engloba o síndrome de Peutz-Jeghers, associado a mutações germinais no gene *STK11* e o síndrome de polipose juvenil, onde se encontram descritas mutações germinais nos genes *SMAD4* e *BMPRIA*, e a polipose serreada. Na polipose serreada podem ocorrer três vias de carcinogénese distintas: a via serreada e a via alternativa em que as lesões precursoras são pólipos serreados ou adenomas com componente viloso e a via tradicional na qual a lesão precursora é o

pólipo adenomatoso (Leggett & Whitehall 2010). No entanto, até à data ainda não se identificou a causa molecular responsável por este síndrome hereditário.

1.3.2. Cancro do cólon e recto hereditário não associado a polipose

O cancro do cólon e recto não associado a polipose apresenta uma transmissão autossómica dominante, uma idade jovem de diagnóstico (<45 anos) e um risco aumentado de desenvolvimento de CCR. Este é o síndrome hereditário de CCR mais frequente representando cerca de 5% de todos os casos de CCR diagnosticados (Lindor et al. 2005).

Clinicamente o HNPCC é definido pelo preenchimento dos critérios de Amesterdão (CA). Inicialmente estes contemplavam apenas a história familiar de CCR (CA-I), tendo sido posteriormente revistos (CA-II) em 1998, de modo a englobarem também os tumores extra-cólicos característicos deste síndrome e desta forma permitir o diagnóstico de um maior número de famílias (Tabela 1) (Umar et al. 2004). Actualmente o HNPCC engloba dois síndromes, o síndrome de Lynch (SL) e o cancro do cólon e recto familiar do tipo X (FCCTX) (Jasperson et al. 2011).

Tabela 1. Critérios de Amesterdão I e II (Umar et al. 2004)

<u>Critérios de Amesterdão I</u>
<ol style="list-style-type: none">1. Pelo menos três indivíduos com cancro do cólon e recto confirmado por histologia, sendo um deles familiar em 1º grau dos outros dois;2. Pelo menos duas gerações sucessivas afectadas;3. Pelo menos um indivíduo com CCR com idade de diagnóstico antes dos 50 anos;4. Exclusão de PAF.
<u>Critérios de Amesterdão II</u>
<ol style="list-style-type: none">1. Pelo menos três indivíduos com CCR ou tumor extra-cólico do espectro de síndrome de Lynch (endométrio, ovário, estômago, intestino delgado, ureter, pélvis renal, tracto hépato-biliar e cérebro) histologicamente confirmado, sendo um deles familiar em 1º grau dos outros dois;2. Pelo menos duas gerações sucessivas afectadas;3. Pelo menos um dos casos de CCR diagnosticado antes dos 50 anos;4. Exclusão de PAF.

1.3.2.1. Síndrome de Lynch

Durantes muitos anos os termos HNPCC e síndrome de Lynch eram utilizados como sinónimos, uma vez que estas famílias preenchiam os CA e apresentavam mutações nos genes do sistema de reparação de erros de DNA do tipo *mismatch* (MMR). No entanto à medida que mais famílias foram sendo estudadas, verificou-se que cerca de metade destas não apresentavam mutações germinais nestes genes (Lynch & de la Chapelle 2003). Assim em 2005 o SL passou a ser definido em base genética pela presença de mutações nos genes MMR, independentemente da história familiar (Boland 2005).

O sistema MMR compreende as proteínas MSH2, MSH3, MLH1, MSH6, PMS1 e PMS2, que permitem a reparação de emparelhamentos incorrectos e pequenas deleções/inserções de nucleótidos durante a replicação de DNA, erros esses, que não são detectados pela polimerase aquando o *proofreading*. Os complexos *MutSa* (MSH2-MSH3) e *MutSβ* (MSH2-MSH6) apresentam afinidade para estes erros no DNA, recrutando em seguida o complexo formado pelas proteínas MLH1 e PMS2 (*MutL*). Em seguida é recrutada uma enzima, a exonuclease, que permite a excisão dos nucleótidos mal emparelhados e por fim a cadeia de DNA é sintetizada de novo pela DNA polimerase sendo reposta a informação genética (Boland & Goel 2011).

Clinicamente o SL caracteriza-se pela presença de CCR numa idade jovem, preferencialmente localizado no cólon proximal, quando comparado com casos esporádicos de CCR. Histologicamente os tumores são pouco diferenciados, com características mucinosas e apresentam infiltrado inflamatório. Os indivíduos com SL apresentam ainda tumores extra-cólicos, sendo comum a presença de tumores no endométrio, estômago, ovário e pélvis renal (Jasperson et al. 2011).

Uma característica molecular dos tumores dos indivíduos de síndrome de Lynch é a presença de instabilidade de microssatélites. As sequências de microssatélite são regiões repetitivas de DNA, polimórficas entre indivíduos mas únicas e uniformes em todas as células do indivíduo. Estas regiões são alvos de *slippage* por parte da DNA polimerase, aquando a replicação do DNA levando a pequenas deleções/inserções de nucleótidos, as quais não são reparadas pelo sistema MMR, por este se encontrar deficiente, levando assim à formação de alelos adicionais nas células tumorais e consequentemente à MSI (Boland & Goel 2011).

Esta característica dos tumores de SL é utilizada para a identificação de novos casos quando as famílias são muito pequenas ou não se tem conhecimento da história familiar, ou seja, quando não preenchem os CA. Nestes casos são utilizados os critérios de Bethesda, definidos em 1997 na conferência do *National Cancer Institute* (Tabela 2) para selecção dos indivíduos em cujos tumores deve ser pesquisada a presença de MSI. Estes critérios são mais abrangentes que os CA pois necessitam apenas do preenchimento de um dos parâmetros. A análise de MSI é realizada utilizando os denominados marcadores de Bethesda, dois marcadores de mononucleótidos (BAT26 e BAT25) e três marcadores de dinucleótidos (D2S123, D5S346 e D17S250). Um tumor é considerado como MSI-Low (MSI-L), se apresentar instabilidade em apenas um dos marcadores, ou MSI-High (MSI-H), se apresentar instabilidade em 2 ou mais marcadores. Se não apresentar instabilidade de microsatélite nos marcadores de Bethesda o tumor é considerado microsatélite estável (MSS) (Boland & Goel 2011).

Os indivíduos com tumores MSI-H são depois encaminhados para a realização de análise mutacional nos genes MMR.

Tabela 2. Critérios de Bethesda (Umar et al. 2004)

<u>Critérios de Bethesda</u>
<p>Preenchimento de pelo menos 1 dos critérios:</p> <ol style="list-style-type: none"> 1. Indivíduo com CCR diagnosticado antes dos 50 anos; 2. Indivíduo com CCR síncronos ou metacrónicos, ou associado com outros tumores do espectro de síndrome de Lynch, independentemente da idade; 3. Indivíduo com CCR com histologia típica de SL (infiltrado linfocitário, reacção <i>Crohn-like</i>, tumores mucinosos ou com diferenciação em “anel de sinete” ou padrão de crescimento medular), diagnosticado antes dos 60 anos; 4. Indivíduo com CCR e 1 ou mais familiares em 1º grau com um tumor do espectro de SL, um dos quais diagnosticado em idade inferior a 50 anos; 5. Indivíduo com CCR e 2 ou mais familiares de 1º ou 2º grau com tumores de espectro de SL, independentemente da idade.

1.4. Cancro do cólon e recto familiar do tipo X

Como dito anteriormente, as famílias HNPCC apresentam alguma heterogeneidade, sendo que em cerca de metade das famílias que preenchem os CA não é detectada uma mutação germinal nos genes MMR e os tumores não apresentam MSI. Deste modo em 2005, o termo cancro do cólon e recto familiar do tipo X foi sugerido para definir estas famílias (Lindor et al. 2005). Por comparação das características clínicas entre as famílias de síndrome de Lynch e as de FCCTX (Tabela 3), conclui-se que estas últimas apresentam uma idade de diagnóstico mais tardia, tumores com localização preferencialmente distal, menor frequência de tumores extra-cólicos, maior número de pólipos mas um risco diminuído para o desenvolvimento de CCR o que sugere uma tumorigénese mais lenta (Lindor 2009).

Tabela 3. Principais características das famílias de síndrome de Lynch e FCCTX (Lindor, 2009)

	Síndrome de Lynch	FCCTX
Idade de diagnóstico	±45 anos	>50 anos
Mutações nos genes MMR	+	-
Tumor	MSI	MSS
Localização	Proximal	Distal
Pólipos	Poucos	Mais
Frequência de tumores extra-cólicos	Elevada	Baixa
Risco de CCR	Elevado	Diminuído
Tumorigénese	Rápida	Lenta

Num estudo realizado pelo grupo de Gastreenterologia da UIPM foi efectuada a caracterização clínica de 15 famílias FCCTX assim como a análise de LOH nos *loci* dos genes *APC*, *TP53*, *SMAD4* e *DCC*, de forma a perceber o papel da via CIN nos tumores FCCTX. Para além disso, foi ainda avaliada a presença de mutações somáticas nos genes *APC* e *KRAS* e o perfil de metilação nos genes *MGMT* e MMR. Os resultados obtidos demonstraram que existe uma maior prevalência de adenomas nos casos de FCCTX quando comparado com o número de carcinomas (66% *versus* 34%, respectivamente) e que a localização preferencial dos carcinomas tende a ser superior no recto, cerca de 48%. As várias análises revelaram duas entidades

molecularmente distintas entre os tumores das famílias FCCTX: os que apresentaram LOH nos *loci* dos genes supressor de tumor (72%), mutações somáticas nos genes *APC* e *KRAS*, bem como metilação no promotor do gene *MGMT*, e os que não apresentaram LOH nos mesmos *loci* (28%), menor frequência de mutações somáticas e praticamente ausência de metilação, sugerindo o envolvimento da via CIN no desenvolvimento de apenas uma parte dos tumores FCCTX (Francisco et al. 2011).

Até à data não se conhece, a causa molecular deste síndrome hereditário sendo, por isso, necessário identificar a sua origem genética.

1.4.1. Mapeamento de *loci* de susceptibilidade para FCCTX

No âmbito do estudo do FCCTX, na mesma instituição foi ainda realizada uma análise de *linkage* utilizando *arrays* de 50.000 SNPs (*single nucleotide polymorphism*) com o objectivo de determinar regiões de *linkage* com a doença em 2 famílias FCCTX, as quais foram seleccionadas, de entre o grupo de 15 famílias, por serem as mais informativas uma vez que se dispunha de material biológico de vários indivíduos, tanto afectados como não afectados. Esta análise permitiu a identificação de três regiões cromossómicas candidatas: 13q, 16q e 21q.

Após análise de *linkage* e LOH utilizando marcadores de microssatélite foi excluída a região 16q, que não revelou *linkage* com a doença, e seleccionadas as regiões 13q e 21q (Belo 2010). Os marcadores de microssatélite utilizados para a análise de LOH permitiram também delimitar a zona de susceptibilidade nestas regiões, nos tumores das famílias FCCTX, tendo sido determinada uma região mínima de LOH de 0,84Mb e de 1,3Mb, para as regiões 13q e 21q respectivamente. No mesmo estudo, verificou-se ainda que as perdas de heterozigotia no cromossoma 13 eram mais frequentes em adenomas, enquanto que as perdas no cromossoma 21 eram mais frequentes em carcinomas, o que sugere que os eventos em 13q ocorrem ao nível da iniciação tumoral enquanto os eventos em 21q ocorrem ao nível da progressão tumoral (Pereira 2013; Belo 2010). Até à data, nem a região 13q nem 21q foram propostas como região de susceptibilidade para o desenvolvimento de CCR. Existe apenas o conhecimento de uma frequência elevada de ganhos na região 13q, abrangendo por vezes todo o braço longo do cromossoma, quer em carcinomas primários quer em CCR metastático (Knösel et al. 2002; Diep et al. 2006). Assim, devido à presença de ganhos cromossómicos frequentes em CCR esporádico, será importante efectuar um estudo de modo a quantificar o número de cópias nesta

região nos tumores FCCTX, de modo a concluir a presença de eventuais amplificações ou deleções cromossómicas. Este estudo permitirá esclarecer os resultados de LOH obtidos em estudos prévios, correspondem a amplificações já descritas noutros estudos em CCR esporádico ou se em alguns casos é detectada deleção, que será sugestiva de um gene supressor tumoral nesta região.

1.4.2. Genes candidatos de susceptibilidade para o FCCTX na região 13q32-33

Dos vários genes contidos na região de interesse foram seleccionados alguns genes candidatos para análise de mutações germinais, de acordo com os seguintes critérios de selecção: genes expressos no cólon; genes que codificam proteínas envolvidas directa ou indirectamente em processos celulares associados à tumorigénese.

Nos parágrafos seguintes, descreve-se de forma sucinta alguma informação relevante sobre os genes candidatos estudados até à data no grupo de Gastreenterologia, no contexto da susceptibilidade para o FCCTX.

O gene *STK24* (*Serine/Threonine Kinase 24*) codifica uma proteína cinase que promove a apoptose em resposta ao *stress* celular por activação da via das caspases. A associação de uma proteína deste tipo como é o caso da STK11 e o síndrome de Peutz-Jeghers tornaram este gene um candidato para estudo em casos de FCCTX (Hennig et al. 2012).

O gene *DOCK9* (*Dedicator of Cytokinesis 9*) codifica uma proteína que actua na troca de GTP por GDP activando especificamente a proteína G Cdc42. A regulação específica por parte da Cdc42 é fundamental para a homeostase celular, encontrando-se a sua desregulação na base de várias patologias, entre elas o cancro (Sinha & Yang 2008; Wheeler et al. 2013).

O gene *METTL21C* (*Methyltransferase Like 21C*), também conhecido como *C13orf39*, codifica uma proteína com função de metiltransferase, semelhante à função desempenhada pela proteína metilguanina metiltransferase (MGMT) a qual desempenha um papel fundamental no cólon na reparação de erros induzidos por *stress* oxidativo. O gene *MGMT* encontra-se frequentemente metilado em carcinomas

esporádicos do cólon, realçando o interesse em estudar o gene *METTL21C* no âmbito do FCCTX (Shen et al. 2005).

O gene *SLC10A2* (*Solute Carrier Family 10 (Sodium/Bile Acid Cotransporter Family), Member 2*), codifica um cotransportador (ASBT) necessário para a reabsorção de ácidos biliares pelas células apicais do íleo. Mutações neste gene são a causa primordial de perturbações na deficiente absorção de ácidos biliares o que pode gerar um processo inflamatório e consequentemente desenvolver um microambiente propício ao desenvolvimento tumoral (Jung et al. 2004).

O gene *TEX30* (*Testis Expressed 30*), também conhecido como *C13orf27* codifica uma proteína com uma função hipotética de hidrolase. Este gene merece ser alvo de estudo pois este tipo de enzima encontra-se associado à progressão tumoral, nomeadamente na regulação da proliferação tumoral (Liu et al. 2012).

O gene *TPP2* (*Tripeptidyl Peptidase II*) é constituído por 29 exões, e encontram-se descritos 8 transcritos alternativos deste gene sendo que apenas dois deles originam um produto funcional. O transcrito alternativo difere do transcrito *wild type* (*wt*) por apresentar um codão de iniciação diferente e pela presença de um exão adicional. Este exão deriva de uma inserção de 39pb codificando 13 aminoácidos na região C-terminal da proteína, no entanto a sua função é desconhecida (Tomkinson et al. 1994). O gene *TPP2* codifica uma aminotripeptidase, a maior protease citosólica eucariota, a qual se encontra descrita como podendo desempenhar várias funções, desde a regulação da síntese e degradação proteica, à regulação da reparação do DNA e da apoptose (Peters et al. 2011). A *TPP2* intervém após a clivagem de proteínas pelo proteossoma, degradando os oligopéptidos em tripéptidos.

Alguns destes genes foram já excluídos pelo grupo de Gastrenterologia, por ausência de alterações relevantes, nomeadamente os genes *STK24* (Pires 2011), *DOCK9* (Zhao 2012) e *METTL21C* (Póvoa 2011). De entre os que não foram ainda excluídos, encontra-se o gene *TPP2* no qual foi identificado um transcrito alternativo, ainda não descrito, que parece apresentar alterações de expressão entre diferentes tumores, sendo necessário esclarecer o seu papel no contexto do FCCTX, nomeadamente através da análise de expressão e da eventual deleção total do gene; o gene *TEX30* para o qual é ainda necessário realizar a análise de grandes deleções/inserções com o intuito de excluir a presença de possíveis deleções/inserções em um ou mais exões e a análise de expressão do transcrito *wt*

deste gene; por fim o gene *SLC10A2* no qual foram identificadas, num estudo prévio, três mutações possivelmente patogénicas (p.V98I, p.P290S e p.A171S), a mutação V98I revelou uma aparente segregação com a doença e as mutações P290S e A171S revelaram, através de análise *in silico* e da descrição na literatura (Wong et al. 1995), carácter possivelmente patogénico (Pereira 2013). Neste estudo será necessário realizar uma análise de segregação da mutação A171S na família onde foi identificada a mutação, de forma a esclarecer se pode contribuir para o fenótipo desta família, em associação uma variante previamente identificada no gene *APC* (p.E1317Q). Relativamente às mutações V98I e P290S que apresentaram uma baixa frequência na série FCCTX e na população europeia, será necessário avaliar a sua frequência na população portuguesa através da análise de indivíduos saudáveis.

1.4.3. Genes candidatos de susceptibilidade para o FCCTX na região 21q11

Para a selecção dos genes candidatos na região 21q11, previamente delimitada, foram utilizados os critérios de selecção descritos no ponto anterior (1.4.2).

Nesta região tinha sido apenas estudado o gene *NR1P1* (*Nuclear Receptor Interacting Protein 1*) que codifica a proteína RIP140, a qual regula a actividade transcricional do receptor de estrogénios. No entanto, a análise mutacional germinal efectuada neste gene em indivíduos FCCTX não revelou a presença de qualquer mutação potencialmente patogénica (Pereira 2013).

Após exclusão de mutações no gene *NR1P1*, foram seleccionados para estudo, nomeadamente análise de mutações pontuais, análise de grandes deleções/inserções e análise de expressão, ao nível germinal, os genes *HSPA13* e *SAMSNI* com base nos critérios de selecção referidos anteriormente e no trabalho de Francisco e colaboradores, o qual revelou dados importantes acerca dos tumores FCCTX, nomeadamente a frequente metilação do gene *MGMT* juntamente com a mutação do gene *KRAS* (Francisco et al. 2011), assim como metilação de alguns genes MMR, num subgrupo de tumores FCCTX. Este padrão de alterações moleculares sugere uma possível condição de susceptibilidade relacionada com o *stress* oxidativo a qual pode estar relacionada com a resposta aos erros no DNA devidos ao *stress* oxidativo/agentes alquilantes, função esta que se relaciona de forma directa ou indirecta com a função dos genes aqui seleccionados.

1.4.3.1. *Microsomal Stress-70 Protein ATPase Core (HSPA13)*

O gene *HSPA13* também conhecido como *STCH* (*Stress 70 Protein Chaperone*), codifica uma proteína que pertence à família das *heat shock protein 70* (Hsp70) e é constituído por 5 exões. Encontram-se descritos 3 transcritos deste gene, sendo que apenas dois deles originam um produto funcional (Figura 2). O transcrito alternativo (Transcrito 2) difere do *wild type* (Transcrito 1) por não ter o exão 3, por apenas parte do exão 4 e o exão 5 serem codificantes e por a região 3'UTR ser menor (Ensembl 2014b).

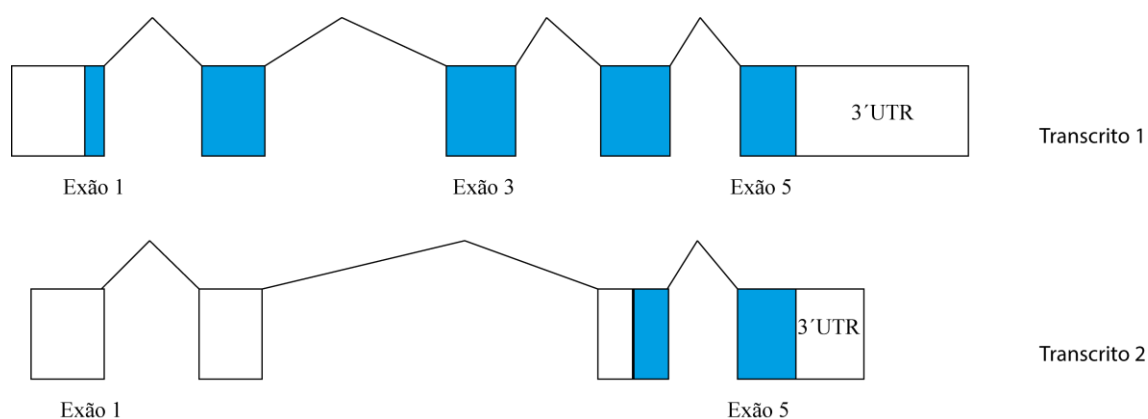


Figura 2. Representação esquemática dos dois transcritos codificantes do gene *HSPA13*. Branco – Região não codificante; Azul – Região codificante; Traço – Intrão.

A proteína codificada por este gene é expressa em níveis mais elevados no plasma sanguíneo, no rim e no fígado (Genecards® 2014b).

As proteínas da família Hsp70 contêm um domínio ATPase na região N-terminal e um domínio de ligação ao substrato na região C-terminal. O domínio ATPase é essencial para a hidrólise de ATP a ADP, permitindo regular a afinidade da enzima para o substrato (Mayer & Bukau 2005). A proteína HSPA13, constituída por 471 aminoácidos (Figura 3) apresenta características distintas das outras proteínas da sua família, nomeadamente uma sequência sinal hidrofóbica na região N-terminal e ausência de um domínio de ligação ao substrato na região C-terminal, sendo ainda regulada pelo ião Ca^{2+} e não pela presença de ATP (Aoki et al. 2005).



Figura 3. Representação esquemática dos domínios conhecidos da proteína HSPA13.

A proteína *HSPA13*, bem como as restantes da sua família, está envolvida em processos celulares de *folding* de proteínas recém sintetizadas ou desnaturadas, encontrando-se assim, descrita como *chaperone* (Mayer & Bukau 2005).

Estudos recentes sugerem que esta proteína pode estar envolvida na apoptose mediada pela via TRAIL (*TNF-related Apoptosis-inducing ligand*) (Yamagata et al. 2008). A via TRAIL (Figura 4) é activada através da ligação do ligando ao receptor TRAIL que recruta a proteína adaptadora *FAS-associated protein with death domain* (FADD) levando à formação de um complexo de indução de morte celular. Este complexo recruta a caspase 8 que oligomeriza e activa-se, de seguida cliva a caspase 3, activando-a, o que conduz ao processo de apoptose celular (Wang & El-Deiry 2003; Smallridge 2003).

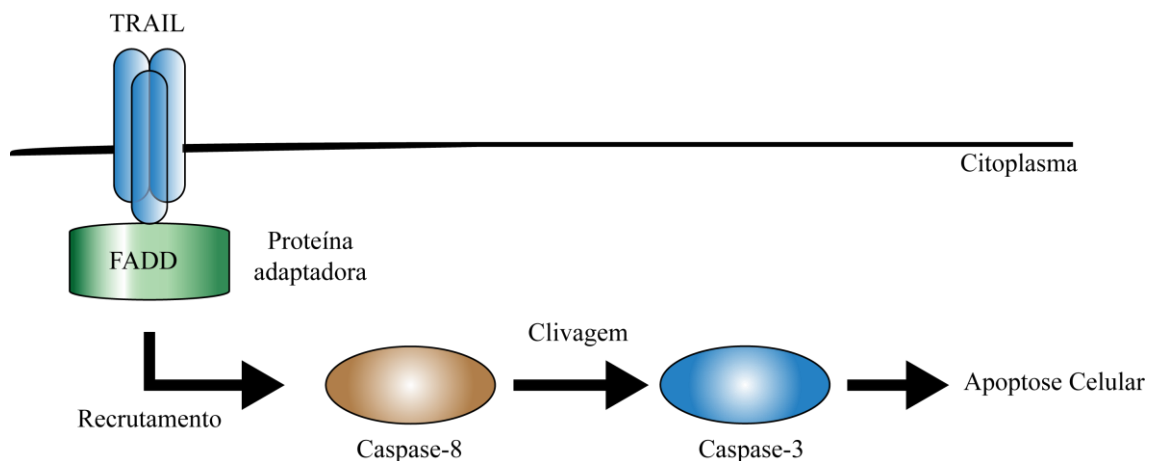


Figura 4. Representação esquemática da via de sinalização pró-apoptótica TRAIL, com representação das suas principais proteínas (adaptado de Smallridge 2003).

A sobreexpressão de proteínas Hsp70 está relacionada com a resistência à apoptose por inibição do apoptossoma sendo sugerida uma possível terapêutica de inibição desta proteína para sensibilizar os tumores à quimioterapia (Sharma &

Masison 2009). No estudo de Yamagata e colaboradores, foi sugerido que a HSPA13 interage com proteínas da sua família de forma a interferir com a actividade anti-apoptótica (Yamagata et al. 2008). Para além deste estudo, Grizenkova e colaboradores, constataram que a sobreexpressão de HSPA13 em células neuronais levava à activação da via TRAIL, sendo que esta é apenas induzida em situação patológica, associada à necrose tumoral (Grizenkova et al. 2012). Estas evidências sugerem uma forte relação entre a apoptose e a HSPA13 e permitem extrapolar que esta proteína desempenha um papel pró-apoptótico em células cancerígenas e que a sua expressão pode estar associada com a resposta à terapêutica e com a agressividade tumoral.

Num estudo de Aoki e colaboradores foi realizada a análise de 66 genes localizados na região 21p11-21q22, utilizando como marcadores genéticos 126 polimorfismos de nucleótido único (SNP) em sangue periférico. Dos SNP analisados 5 deles foram detectados no gene *HSPA13* e apresentam uma associação com o carcinoma gástrico não localizado no cárdia. Este tipo de carcinoma não localizado no cárdia está associado com a infecção por *Helicobacter pylori*, sugerindo que pode existir uma relação entre a infecção por *H. pylori* e uma possível susceptibilidade conferida pelos SNP para o desenvolvimento deste tipo de carcinoma (Aoki et al. 2005).

1.4.3.2. *SAM Domain, SH3 Domain And Nuclear Localization Signals 1 (SAMSNI)*

O gene *SAMSNI* também conhecido como *HACSI* (*Hematopoietic Adapter-Containing SH3 and Sterile α -motif (SAM) domains 1*), é constituído por 9 exões dos quais 8 são codificantes. Encontram-se descritos 5 transcritos deste gene, dos quais, apenas 3 codificam um produto funcional (Figura 5). Os transcritos alternativos diferem do *wt* (transcrito 1) por possuírem um o exão 1 diferente e no caso do transcrito 3 ainda pela ausência dos exões 3-6 (Ensembl 2014c).

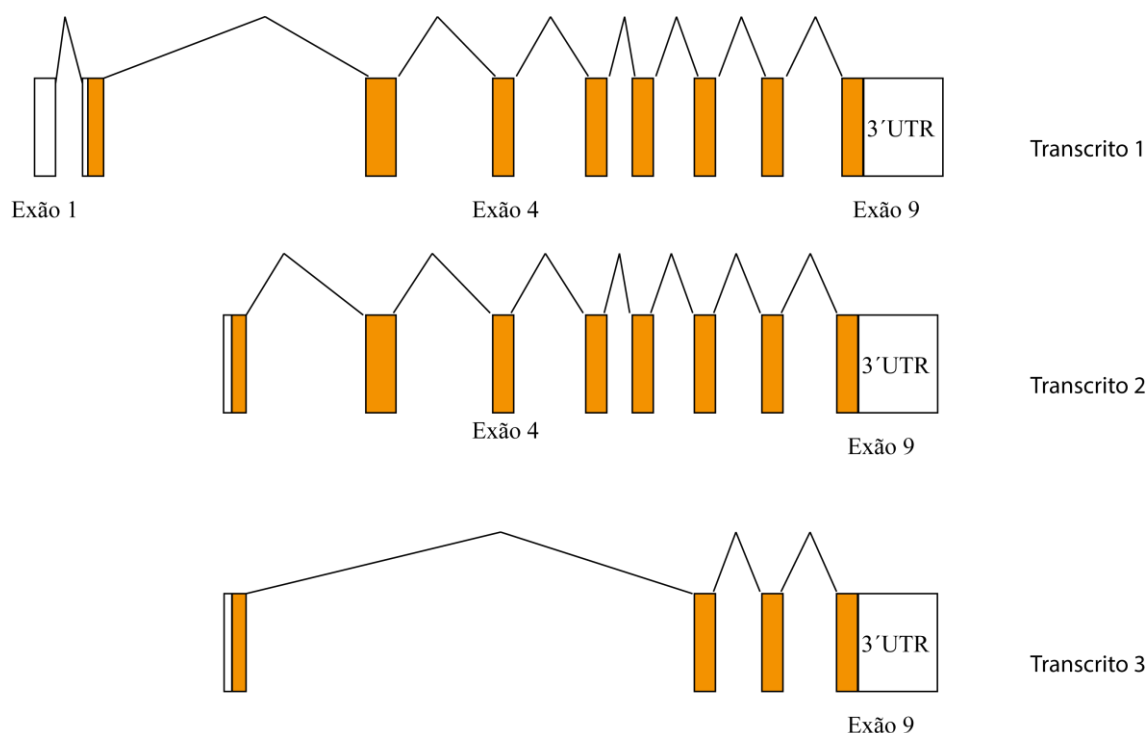


Figura 5. Representação esquemática dos transcritos codificantes do gene *SAMSNI*. Branco – Região não codificante; Laranja – Região codificante; Traço – Intrão.

O gene *SAMSNI* localiza-se numa região do cromossoma 21 associada com patologias hematopoiéticas devido a translocações frequentes nesta região (Holleben et al. 2011). A proteína codificada por este gene é predominantemente expressa no plasma sanguíneo, no coração, no rim e no fígado (Genecards® 2014c). Esta proteína tem como principal função a inibição do sistema imunitário através da via inibidora PIR-B em associação com a proteína tirosina cinase Lck/Yes-related novel (Lyn). A proteína Lyn desempenha a função de regulação da resposta imunitária, nomeadamente ao nível dos linfócitos B (Wang et al. 2010).

A proteína SAMSNI é constituída por 441 aminoácidos, contém um domínio de interacção proteica, um domínio de fosforilação tirosina cinase, um domínio *SRC Homology 3* (SH3) e um domínio *Sterile α* (SAM) (Figura 6). Os domínios SH3 são normalmente associados a proteínas adaptadoras, sugerindo uma função semelhante desta proteína (Wang et al. 2010). A SAMSNI contém ainda 3 sequências de localização nuclear, sugerindo uma função ao nível do núcleo (Claudio et al. 2001).

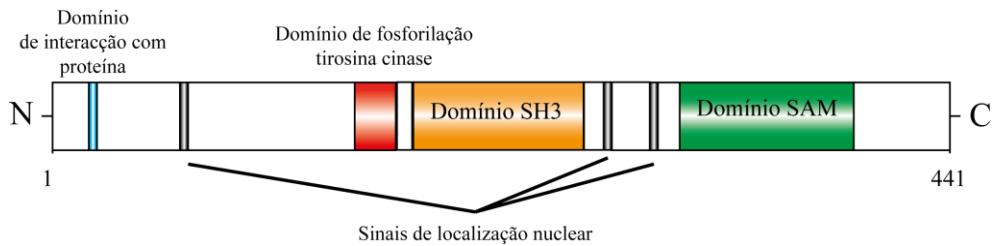


Figura 6. Representação dos domínios conhecidos da proteína SAMSNI.

Poucos estudos relacionam esta proteína com a tumorigénese, no entanto Yamada e colaboradores, efectuaram um *screening* utilizando 12 marcadores de *sequence tagged-site* (STS) em 85 linhas celulares de carcinoma do pulmão, resultando na identificação de uma deleção homozigótica em 21q21.1 que posteriormente foi limitada a 3,4Mb, onde estão localizados os genes *HSPA13*, *NRIP1* e *SAMSNI*, incluídos neste estudo ou estudados previamente. Foi ainda realizada uma análise de expressão de dois genes contidos nesta região, um deles *SAMSNI*, que revelou uma diminuição de expressão em todas as linhas celulares de carcinoma do pulmão utilizadas. Com base neste estudo, Yamada sugere que o gene *SAMSNI* pode actuar como gene supressor de tumor em cancro do pulmão (Yamada et al. 2008). Noutro estudo, Watanabe e colaboradores, identificaram 20 genes que permitiram diferenciar os casos de colite ulcerosa que desenvolveram carcinoma e os que não desenvolveram. Entre estes 20 genes inclui-se o gene *SAMSNI*, sugerindo que este pode estar envolvido na progressão tumoral em casos de colite ulcerosa (Watanabe et al. 2011).

1.4.4. *Bone morphogenetic protein receptor, type IA (BMPRI1A)*

O gene *BMPRI1A* também conhecido como *activin-like kinase 3 (ALK3)* localiza-se no cromossoma 10 (10q22.3) e codifica para um receptor serina-treonina cinase tipo I que tem como ligando a proteína *bone morphogenetic protein* (BMP). Este gene é constituído por 13 exões dos quais 11 são codificantes. Encontram-se descritos 2 transcritos deste gene, no entanto, apenas um codifica para um produto funcional (Ensembl 2014a). A proteína codificada por este gene é expressa em níveis mais elevados no plasma sanguíneo e no rim (Genecards® 2014a).

Esta proteína é constituída por 532 aminoácidos e encontra-se localizada na membrana celular. Quando o ligando BMP, pertencente à família *Transforming*

growth factor (TGF- β), se liga ao receptor conduz à formação de um complexo entre o receptor tipo I e o tipo II. O receptor tipo II fosforila o receptor tipo I, que por sua vez fosforila as proteínas SMAD1/5/8, activando-as, as quais formam um complexo com a proteína SMAD4 (Figura 7). Este complexo é então importado para o núcleo onde actua como factor de transcrição regulando genes de proliferação e diferenciação celular, esta via actua como supressora do crescimento celular (Cao & Chen 2009; Attisano & Wrana 2002).

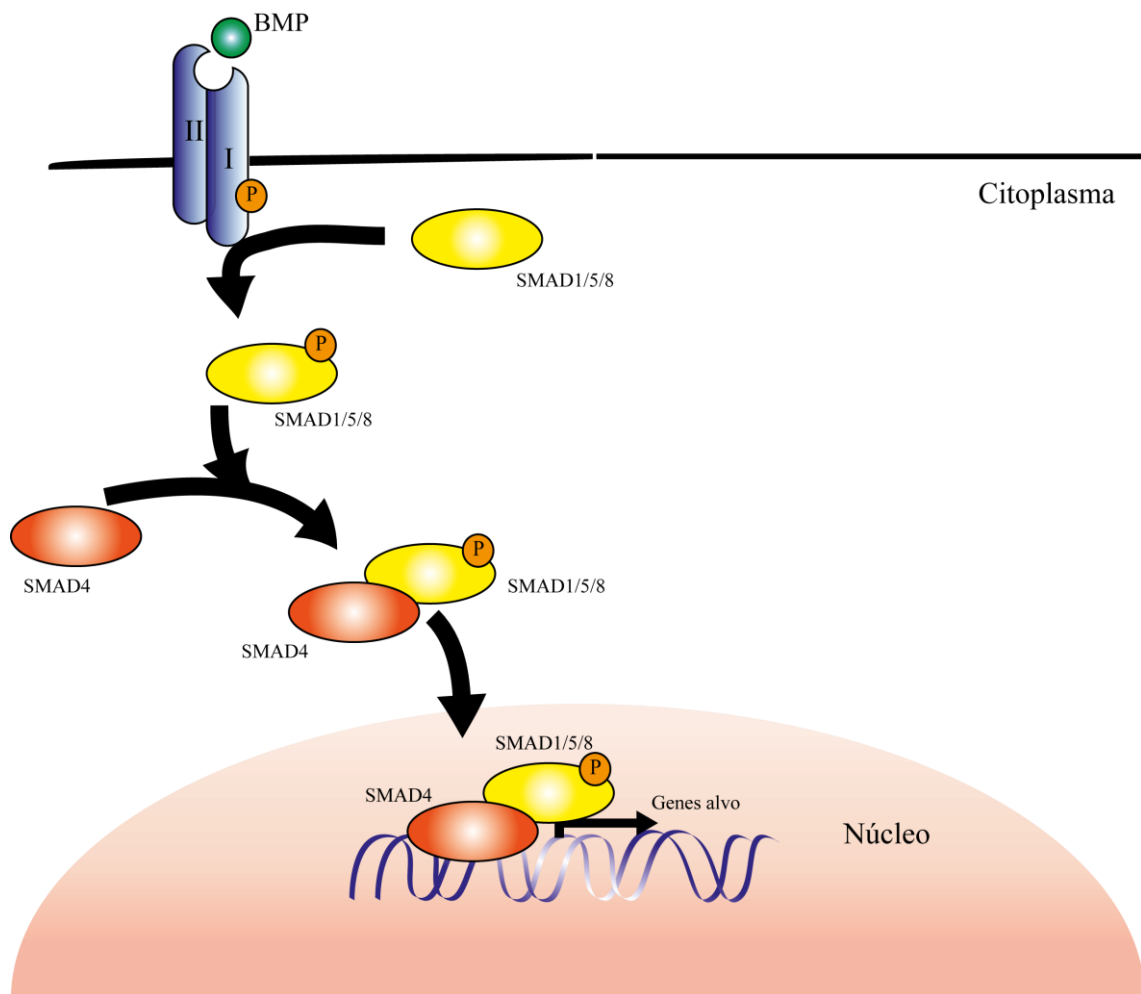


Figura 7. Representação esquemática da via de sinalização BMP com o ligando BMP e as principais proteínas da via (adaptado de Hardwick et al. 2008); P – Proteína fosforilada.

Hardwick e colaboradores estudaram um dos ligandos da via BMP, a proteína BMP2, que actua de forma autocrina nas células maduras do epitélio cólico, constatando assim que esta via promove a supressão do crescimento celular e a apoptose levando à renovação do epitélio (Hardwick et al. 2004). O desenvolvimento

de CCR provém da ausência de regulação do crescimento celular e de apoptose. A via do TGF- β é uma das principais vias que sofre desregulação, nomeadamente em 80% dos casos de CCR. A via de sinalização BMP, que apresenta semelhanças com a via do TGF- β , pode ser inactiva por mutações germinais no gene *BMPRIA* levando à predisposição para o síndrome de polipose juvenil. Segundo o estudo de Beck e colaboradores realizado em linhas celulares de CCR esporádico, a via BMP parece permanecer intacta ao contrário da via do TGF- β . No entanto em casos de síndrome de polipose juvenil a inactivação da via BMP predispõe a um risco aumentado de desenvolvimento de CCR (Beck et al. 2006).

Recentemente, no trabalho de Nieminen e colaboradores, foi efectuado um estudo de *genome-wide linkage* para uma família que preenchia os CA-I mas que não apresentava nenhuma mutação patogénica nos genes MMR, ou seja, com características de FCCTX. Foi observado *linkage* para a região 10q23, a qual contém os genes *PTEN*, *MINPP1* e *BMPRIA*, e após análise mutacional do gene *BMPRIA* foi identificada uma mutação patogénica, p.Glu88del, que segrega com a doença. Em seguida foram analisadas as restantes famílias incluídas neste estudo, tendo sido identificada numa família uma mutação num local de *splicing*, c.68-10_68+14del, a qual leva ao *skipping* do exão 2 do gene *BMPRIA* (Nieminen et al. 2011). Os resultados deste trabalho demonstram que alterações no *BMPRIA* não estão restritas a famílias com polipose juvenil, e que este gene pode ser um possível candidato na susceptibilidade para o FCCTX.

2. Objectivos

O presente trabalho teve como objectivos a avaliação do envolvimento de genes específicos na susceptibilidade para o FCCTX, nomeadamente do gene *BMPRI1A* e de dois genes localizados na região cromossómica 21q11 (nomeadamente *HSPA13* e *SAMSN1*), previamente identificada como uma possível região de susceptibilidade para o FCCTX. Pretendeu-se, contribuir para o esclarecimento da base molecular deste síndrome, através da análise mutacional dos genes acima mencionados, nomeadamente análise de mutações pontuais, análise de grandes deleções e inserções e análise de expressão.

Pretendeu-se, ainda, concluir a avaliação do envolvimento na susceptibilidade para o FCCTX de dois genes localizados em 13q32-q33 e analisados em estudos prévios, através do estudo da possível patogenicidade de três mutações previamente identificadas no gene *SLC10A2*, assim como do estudo dos genes *TPP2* e *TEX30* ao nível da expressão do RNA.

Pretendeu-se ainda esclarecer os mecanismos moleculares envolvidos nas perdas de heterozigotia detectadas frequentemente na região 13q em tumores FCCTX, descritas num estudo prévio, através da análise do número de cópias por PCR quantitativo.

3. Materiais e métodos

3.1. Materiais

Para a realização deste estudo seleccionaram-se 15 famílias que cumprem os requisitos para FCCTX, seguidas na Clínica de Risco Familiar do Cancro do Cólon e Recto do IPOLFG, EPE, ou seja, famílias que preenchem os CA, não apresentam mutações nos genes MMR e cujos tumores não possuem instabilidade de microssatélites.

Para a análise mutacional dos genes *HSPA13*, *BMPRIA* e *SAMSN1* foi utilizado DNA genómico extraído a partir de sangue periférico de 15 indivíduos índice das famílias FCCTX e 2 familiares de uma das famílias (L55), família que apresenta *linkage* com a doença nas regiões cromossómicas 13q32-33 e 21q11. Estes 2 familiares foram seleccionados de forma a aumentar a probabilidade de encontrar uma mutação que esteja associada com a doença, sendo que o *linkage* nas duas regiões não exclui a possibilidade de alguns indivíduos possuírem alterações apenas em genes localizados na região 13q32-33 e outros nos genes em 21q11.

Na pesquisa de transcritos alternativos dos genes *TEX30* e *SAMSN1* foi utilizado cDNA sintetizado a partir de RNA extraído de sangue periférico de 8 indivíduos de 3 famílias FCCTX, L55, L7 e L63, (3 indivíduos índice e 6 familiares da família L55). Para a conclusão da análise dos transcritos alternativos detectados em estudos anteriores nos fragmentos 5 (inserção de 72pb) e 8 (inserção de 39pb) do gene *TPP2* e do transcrito alternativo encontrado neste estudo do gene *TEX30*, foi utilizado cDNA sintetizado a partir de RNA extraído de sangue periférico de 19 indivíduos com SL e 15 indivíduos saudáveis que constituem o grupo controlo (N).

Relativamente ao estudo da expressão do transcrito *wild type* dos genes, *TEX30*, *SAMSN1*, *HSPA13* e *TPP2* e do transcrito com inserção de 72pb do gene *TPP2*, foi utilizado cDNA sintetizado a partir de RNA extraído de sangue periférico de 7 indivíduos de 2 famílias FCCTX, L55 e L7, (2 indivíduos índice e 5 familiares da família L55). Para o transcrito *wild type* dos genes *SAMSN1* e *HSPA13* foram ainda incluídos 5 indivíduos do grupo N e 2 indivíduos do grupo SL. Para o transcrito *wild type* e para o transcrito com inserção de 72pb do gene *TPP2* foram também incluídos 17 indivíduos do grupo N e 13 indivíduos com SL.

Para o estudos de variabilidade de *copy number* dos genes *SAMSNI*, *HSPA13* e *TPP2* foi utilizado DNA genómico extraído a partir de sangue periférico de 7 indivíduos de 2 famílias FCCTX (2 indivíduos índice e 5 familiares da família L55).

Relativamente ao estudo de variabilidade de *copy number* na região 13q32-33, onde foram analisados os genes *TPP2* e *DOCK9*, foi utilizado DNA extraído a partir de tecido incluído em parafina (FFPE) de 14 tumores FCCTX e 12 mucosas normais, sendo que para dois dos tumores não se possui material biológico de mucosa normal incluído em parafina.

Por forma a concluir o estudo da eventual patogenicidade de três mutações detectadas anteriormente no gene *SLC10A2* foi realizada pesquisa das mutações V98I (exão1) e P290S (exão 5) em DNA genómico extraído a partir de sangue periférico de 48 indivíduos saudáveis de modo a avaliar a frequência destas mutações na população portuguesa em geral. Em relação à mutação A171S (exão 3), foi efectuada uma análise de *linkage* da mutação com a doença, utilizando DNA genómico extraído a partir de sangue periférico de 10 indivíduos da família FCCTX (1 indivíduo índice e 9 familiares da família L84).

3.2.Métodos

3.2.1. Extracção de ácidos nucleicos

As amostras de DNA e RNA utilizadas foram extraídas previamente a este estudo.

3.2.1.1. Extracção de DNA a partir de sangue periférico

A extracção de DNA a partir de sangue periférico, foi efectuada utilizando o kit CITOGENE – Genomic DNA Purification Kit (Citomed). De seguida o DNA foi quantificado num espectrofotómetro Nanodrop 2000 (*Thermo Scientific*) por forma a se efectuar uma diluição de trabalho de 80ng/µl. O DNA *stock* foi armazenado a -20°C.

3.2.1.2. Extracção de DNA a partir de tecido tumoral e mucosa normal incluído em parafina

Para extracção de DNA genómico a partir de tecido incluído em parafina foi utilizado o método de fenol-clorofórmio. O DNA resultante foi quantificado num gel de agarose a 0,8% (p/v) por comparação com amostras de DNA referência de concentrações conhecidas.

3.2.1.3. Extracção de RNA a partir de sangue periférico

O RNA foi extraído utilizando o método descrito para o reagente Trizol (*Invitrogen*) e quantificado num espectrofotómetro Nanodrop 2000 (*Thermo Scientific*). No entanto como o DNA e o RNA absorvem ao mesmo comprimento de onda (260nm), foi necessário efectuar uma electroforese em gel de agarose de 0,8% (p/v), de modo a se observar a integridade do RNA e também possíveis contaminações com DNA.

3.2.2. Amplificação do DNA por reacção em cadeia da polimerase (PCR)

Esta técnica baseia-se na capacidade da DNA polimerase conseguir sintetizar novas cadeias de DNA, *in vitro*, a partir de uma cadeia molde. Comporta três etapas, desnaturação térmica a 94°C, *annealing* dos *primers* e a extensão da nova cadeia de

DNA. As três etapas são repetidas entre 30 a 40 vezes, constituindo o número de ciclos da reacção de PCR, e assim amplificando de forma exponencial o DNA alvo. Para que a amplificação ocorra é necessário um par de *primers* (*forward* e *reverse*), que delimite a zona de interesse no gene alvo; um DNA molde; uma DNA polimerase, sendo a mais usada a Taq polimerase proveniente de *Thermis aquaticus* (pela sua capacidade de ser termorresistente); os desoxinucleótidos (dNTPs) utilizados para a síntese da nova cadeia de DNA; um tampão salino para manter o pH óptimo e constante durante a reacção de amplificação, e uma solução contendo iões magnésio o qual actua como um cofactor essencial para a acção da polimerase.

3.2.2.1. Optimização das condições de PCR

Inicialmente é necessário proceder a uma optimização das condições de amplificação de modo a obter o produto desejado de forma específica. As condições que devem ser aferidas são, a temperatura de *annealing* dos *primers*, a concentração do ião Mg^{2+} , a concentração de *primer*, e o número de ciclos da reacção de PCR. Relativamente à temperatura de *annealing* dos *primers* esta deve ser aumentada sempre que se verifique inespecificidade do produto de PCR ou diminuída no caso em que não se observe a banda de interesse. A concentração de cloreto de magnésio deve ser ajustada, de modo a se obter uma banda específica e cuja quantidade de produto amplificado seja tal que permita efectuar todos os passos posteriores do procedimento experimental. A variação da concentração de *primer* só é efectuada se a variação da temperatura de *annealing* não se demonstrar eficiente na redução de dímeros de *primers* ou na obtenção de um produto específico de DNA. No entanto, estas alterações podem não ser suficientes para se obter um produto de PCR específico, nestes casos, pode ser necessário alterar o *kit* da DNA polimerase, e utilizar por exemplo, uma polimerase com *hotstart* térmico, o que permite obter uma diminuição na inespecificidade do PCR, uma vez que a enzima apenas tem actividade a uma determinada temperatura óptima.

Todas as amostras de DNA utilizadas neste estudo, foram amplificadas para um volume final de reacção de 12,5µl, constituída por 1µl de DNA molde (80ng/µl) e o restante da mistura reaccional, contendo todos os componentes necessários à sua amplificação, referidos anteriormente.

3.2.2.2. Análise electroforética do produto amplificado por PCR em gel de agarose

A electroforese consiste na separação de moléculas, através da aplicação de um campo eléctrico, numa matriz, que funciona como um filtro. A agarose permite que as moléculas de DNA sejam separadas consoante o seu tamanho molecular ao migrarem do pólo negativo para o pólo positivo, sendo que as moléculas mais pequenas migram mais rapidamente enquanto que as maiores migram mais lentamente. Para a visualização do produto de PCR é necessário a utilização de corantes fluorescentes que se intercalam na cadeia dupla de DNA, como por exemplo o brometo de etídio, e que quando expostos à luz ultra-violeta emitem fluorescência.

Todas as soluções utilizadas na análise electroforética e na preparação dos géis de agarose estão descritas no anexo V.

Aos produtos de PCR obtidos da optimização das condições de amplificação são adicionados 3µl de tampão de deposição – *Orange G 5x*, com a finalidade de conferir densidade à amostra de modo a que esta se deposite no fundo do poço de agarose e ainda permitir a visualização da frente de migração, durante a electroforese. Após a deposição das amostras, em gel de agarose 1,2%-2,0% (p/v) (de acordo com o tamanho molecular das amostras) estas foram submetidas a uma voltagem constante de 140V durante um tempo variável entre os 30 minutos e 1 hora e 15 minutos. Juntamente com as amostras foi aplicado no gel um marcador de peso molecular adequado com um padrão de bandas de tamanhos moleculares conhecidos, com a finalidade de confirmar o tamanho do produto amplificado. Após a electroforese em tampão TBE 1x, o gel foi exposto a radiação ultra-violeta, num transiluminador ChemiDoc™XRS+ System (Bio-Rad) com captação da imagem.

3.2.3. Análise mutacional dos genes *HSPA13*, *SAMSN1* e *BMPRIA*

3.2.3.1. Amplificação por PCR dos exões dos genes *HSPA13*, *SAMSN1* e *BMPRIA*

Os *primers* para a amplificação por PCR destes três genes foram desenhados utilizando o *software* NetPrimer (*PREMIER Biosoft*). Para o gene *BMPRIA* foram utilizados 11 pares de *primers* já desenhados anteriormente a este estudo, com o intuito de amplificar os 11 exões codificantes deste gene bem como a sua região 3'UTR. Para o gene *HSPA13* foram desenhados 5 pares de *primers* para amplificação dos 5 exões deste gene e da região 3'UTR, e para o *SAMSN1* foram desenhados 8 pares de *primers* para amplificação dos 8 exões codificantes deste gene bem como da sua região 3'UTR.

Os *primers* utilizados na amplificação dos vários exões encontram-se representados esquematicamente na Figura 8, para o *BMPRIA* (A), *HSPA13* (B) e *SAMSN1* (C), respectivamente. Para a maioria dos casos, os *primers* estão situados nas regiões flangeadoras dos exões de modo a abranger todo o exão e as regiões de *splicing* adjacentes. Para o exão 5 do gene *HSPA13* foi necessário desenhar dois pares de *primers* sobreponíveis, devido ao tamanho deste exão (sendo que a região terminal do produto amplificado pelo primeiro par de *primers* é sobreponível com a região inicial do produto amplificado pelo segundo par), e desta forma conseguir amplificar todo o exão 5 e respectiva região 3'UTR. Relativamente aos outros dois genes, o exão 1 do gene *SAMSN1* e os exões 1 e 2 do gene *BMPRIA* não foram analisados por se tratarem de regiões não codificantes.

As condições para a amplificação dos exões dos genes *BMPRIA*, *HSPA13* e *SAMSN1* assim como os programas de PCR utilizados encontram-se descritos no anexo II.

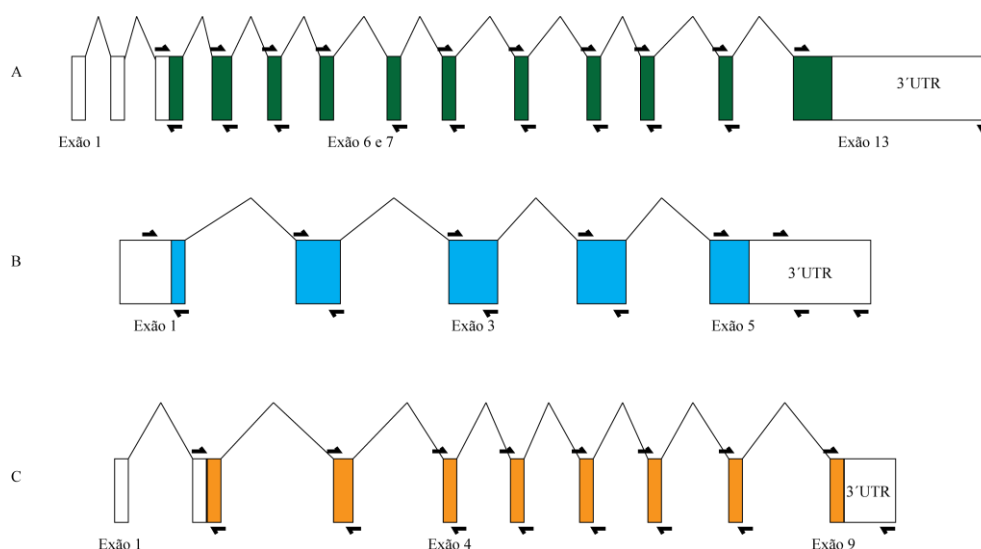


Figura 8. Representação esquemática do gene *BMPRIA* (A), *HSPA13* (B) e *SAMSNI* (C) e respectiva localização dos *primers* (setas) utilizados. Caixas preenchidas – região codificante (exão); Caixas a branco – região não codificante; Linhas – região intrónica.

3.2.3.2. Controlo da eficiência do produto amplificado por PCR em gel de agarose a 2% (p/v)

Os produtos de PCR amplificados foram controlados em gel de agarose de forma a verificar a eficiência da reacção, para tal foi utilizado 3µl do produto amplificado ao qual se adicionou 7µl de tampão de deposição – *Orange 1x* (diluído a partir do tampão de deposição *Orange 5x*). A electroforese decorreu como descrito em 3.2.2.2, tendo-se utilizado o marcador de peso molecular GeneRuler 50bp DNA Ladder (*Thermo Scientific*). No caso do exão 13 do *BMPRIA* foi utilizado o marcador de peso molecular 1kb DNA Ladder (*Promega*), uma vez que o produto esperado era superior a 1000pb.

3.2.3.3. Purificação dos produtos de amplificação por PCR

Posteriormente ao controlo da eficiência do produto de PCR em gel de agarose foi efectuada a purificação do produto amplificado. Foram utilizados dois métodos de purificação diferentes, um método enzimático sempre que se obteve um produto de PCR específico, ou em coluna quando não foi possível eliminar bandas inespecíficas no produto de PCR.

3.2.3.3.1. Purificação enzimática dos produtos de amplificação por PCR

Foi efectuada uma mistura reaccional com as enzimas, *FastAP* 1U/μl e *ExoI* 20U/μl (*Thermo Scientific*), as quais tem a função de eliminar todos os reagentes em excesso na reacção e purificar o DNA amplificado, numa proporção de 2:1, respectivamente. Em seguida, foram distribuídos 1,5μl da mistura em cada tubo contendo o produto de PCR. A purificação foi realizada num termociclador, durante 15 minutos a 37°C, temperatura óptima de actuação das enzimas, seguidos de 15 minutos a 85°C para inactivação das mesmas.

3.2.3.3.2. Purificação em coluna dos produtos de amplificação por PCR utilizando o kit QIAquick® Gel Extraction (*Qiagen*)

Nos casos em que temos mais do que um produto amplificado foi necessário isolar a banda de interesse do gel de agarose. Para tal, o gel foi exposto à radiação ultra-violeta e a banda de interesse excisada com a ajuda de um bisturi e transferida para um tubo tipo *ependorf* de 1,5ml. Posteriormente procedeu-se à purificação do fragmento utilizando o kit QIAquick® Gel Extraction (*Qiagen*), o qual se baseia num procedimento de ligação-lavagem-eluição utilizando uma coluna com membrana de nitrocelulose. Resumidamente, foi efectuada a solubilização do gel de agarose por adição de uma solução tampão ao tubo contendo a agarose excisada, o qual é colocado num banho seco à temperatura de 50°C. De seguida, foi adicionado isopropanol que ajuda na precipitação do DNA, e o volume total foi transferido para uma coluna com tubo de recolha. Esta foi centrifugada durante 1 minuto, de modo a permitir a eluição da solução e que o DNA se ligue à membrana. De seguida, foi efectuada a lavagem do DNA por adição de um tampão que contem etanol, o qual permite a remoção de sais e por fim o DNA foi eluído em tampão TE para um novo tubo tipo *ependorf* de 1,5ml. O protocolo detalhado encontra-se descrito no anexo VI.

De forma a verificar a eficiência da purificação, o produto obtido foi controlado em gel de agarose 1,2% (p/v) como descrito em 3.2.2.2, mas utilizando 2μl de DNA purificado ao qual se adicionou 8μl de *Orange G 1x*. Foi utilizado o marcador de peso molecular GeneRuler 50bp DNA Ladder (*Thermo Scientific*) ou 1kb DNA Ladder (*Promega*) dependendo do tamanho do produto esperado.

3.2.3.4. Sequenciação automática

A sequenciação automática baseia-se nos princípios da replicação do DNA, sendo um método desenvolvido por Sanger e denominado por *dideoxy sequencing* (Applied Biosystems 2009). Este método permite determinar cada nucleótido da sequência de DNA amplificada na reacção de PCR, através da propriedade da DNA polimerase incorporar análogos de nucleótidos com ausência do grupo 3'-hidroxilo, os dideoxirribonucleótidos (ddNTPs). A ausência deste grupo impede a ligação de um nucleótido adjacente, o que leva à terminação da síntese da molécula de DNA e à formação de vários produtos com um nucleótido de diferença.

Cada ddNTP é marcado com um fluorocromo de cor diferente, permitindo a emissão de fluorescência quando excitado pelo laser durante a electroforese capilar, o que leva à determinação dos vários nucleótidos presentes na sequência de DNA em estudo.

3.2.3.4.1. Reacção de sequenciação

A reacção de sequenciação foi efectuada para um volume final de 20µl. Para tal preparou-se uma mistura reaccional contendo 2µl de BigDye®Terminator v1.1, v3.1 5x Sequencing Buffer (*Applied Biosystems*) para manter o pH constante; 2µl do *primer* (*forward* ou *reverse*) numa concentração de 1,6pmol/µl para o alvo que se pretende sequenciar e BigDye™ Terminator v1.1 Cycle (*Applied Biosystems*) que contém uma DNA polimerase, Mg²⁺, dNTPs e ddNTPs. A quantidade de produto de PCR purificado utilizado na reacção, é determinada com base no tamanho molecular do fragmento amplificado e na intensidade da banda obtida após visualização no gel em agarose como descrito em 3.2.2.2. Por outro lado, a quantidade de BigDye™ Terminator v1.1Cycle depende apenas do tamanho do produto amplificado por PCR. Por fim utiliza-se a água bidestilada para perfazer o volume final da reacção.

A reacção de sequenciação foi realizada num termociclador, Veriti, (*Applied Biosystems*) utilizando o programa que se encontra descrito na Tabela 4. A temperatura de *annealing* utilizada na reacção de sequenciação foi a mesma da reacção de PCR excepto nos casos em que a temperatura optimizada para a amplificação dos fragmentos dos diversos genes excedia os 60°C. Nestes casos, utilizou-se a temperatura de 59,5°C, uma vez que, como a temperatura de extensão da enzima é de 60°C, esta poderia ocorrer em simultâneo com o *annealing* dos *primers* e

assim tornar a reacção pouco eficiente. Os produtos obtidos pela reacção de sequenciação foram guardados a 4°C por um período máximo de 24h, até se proceder à sua purificação e precipitação.

Tabela 4. Programa utilizado na reacção de sequenciação

Etapa da reacção	Temperatura (°C)	Duração	Ciclos
Desnaturação inicial	96	5min	1
Desnaturação	95	10seg	25
<i>Annealing</i>	Variável	5seg	
Extensão	60	4min	
Pausa	4	∞	1

3.2.3.4.2. Purificação e precipitação do DNA após a reacção de sequenciação

O método utilizado para a purificação do DNA após a reacção de sequenciação tem como base uma reacção de *salting out*, descrita para o kit BigDye™ Terminator v1.1 Cycle Sequencing (*Applied Biosystems*). Este método utiliza o acetato de sódio que neutraliza a carga dos ácidos nucleicos e ajuda a precipitar o DNA juntamente com o etanol absoluto, e um agente quelante, o EDTA que vai quelatar o Mg^{2+} residual da reacção de sequenciação. Assim foi efectuada uma mistura reaccional contendo de 2µl de EDTA (125mM, pH = 8,0), 2 µl de acetato de sódio (3M, pH = 5), e 50 µl de etanol absoluto num tubo tipo *eppendorf* de 1,5ml. O volume total do produto da reacção de sequenciação foi transferido para o tubo contendo a mistura e homogeneizado em *vortex*. Em seguida, os tubos foram incubados durante 15 minutos à temperatura ambiente e posteriormente centrifugados durante 30 minutos a 14000rpm a 4°C. Após remoção do sobrenadante foi adicionado etanol a 70% (v/v) para lavagem do *pellet* e remoção de sal residual, e nova centrifugação durante 15 minutos. Após a centrifugação o sobrenadante foi descartado e o *pellet* de DNA foi seco num banho seco a 37°C, de modo a remover o

etanol residual que pode interferir com a electroforese capilar. O *pellet* de DNA foi armazenado a 4°C até se proceder à electroforese capilar (Anexo VII).

3.2.3.4.3. Electroforese capilar

A electroforese capilar permite a separação dos produtos da reacção de sequenciação de acordo com o tamanho molecular e com uma resolução de 1 nucleótido. Perto do final da migração no polímero ocorre a excitação do fluorocromo dos ddNTPs por um laser, sendo a emissão de luz detectada por um aparelho óptico e o sinal de fluorescência convertido em dados digitais por um *software*.

3.2.3.4.3.1. Preparação das amostras para electroforese capilar

A preparação das amostras para electroforese capilar consistiu na adição de 17µl de formamida HI-DI (*Applied Biosystems*) a cada tubo contendo o *pellet* seco, seguindo-se uma homogeneização em *vortex* e transferência do volume total para tubos de sequenciação. Em seguida, efectuou-se uma desnaturação térmica das amostras, num termociclador durante 5 minutos a 95°C. Posteriormente, os tubos foram incubados em gelo durante 2 a 3 minutos, e por fim, transferidos para o sequenciador automático ABI Prism™ 310 Genetic Analyser (*Applied Biosystems*) onde se procedeu à electroforese capilar.

3.2.3.4.3.2. Análise de resultados

Os resultados foram obtidos sob a forma de um electroforetograma, após análise com o *software* Sequencing Analysis 3.4.1 (*Applied Biosystems*).

A sequência obtida foi comparada com a sequência normal do gene alvo, retirada da base de dados NCBI, de modo a identificar possíveis alterações.

3.2.4. Análise de grandes deleções/inserções nos genes *TPP2*, *TEX30* e *SAMSN1*

Os genes seleccionados para análise de grandes deleções/inserções foram analisados para mutações germinais pontuais, no entanto através de sequenciação automática não é possível constatar deleções/inserções de 1 ou mais exões. Esta análise foi efectuada utilizando mRNA de modo a se amplificar a totalidade do gene

numa única reacção. Os genes *TPP2* e *TEX30* já tinham sido analisados para mutações pontuais em trabalhos anteriores (Póvoa 2011; Pereira 2013).

3.2.4.1. Reacção de transcrição reversa seguida de PCR (RT-PCR) para síntese de cDNA

A reacção de RT-PCR permite a obtenção do cDNA a partir de um molde de RNA o qual é convertido em cDNA de cadeia simples, formando um híbrido RNA-cDNA e posteriormente amplificado numa molécula de cDNA de cadeia dupla. O RNA utilizado para a reacção foi extraído e quantificado como descrito em 3.2.1.3. Inicialmente foi efectuada uma mistura reacional para um volume final de 7,75µl com um volume de RNA equivalente a 1µg (sempre que possível), água bidestilada tratada com dietilpirocarbonato (DEPC)(*Merck*) que actua como uma DNase, e 0,5µl de hexâmeros (3µg/µl, *Roche*), que funcionam como *primers* aleatórios para o início da síntese de cDNA. Os tubos contendo esta mistura foram incubados a 70°C durante 10 minutos num termociclador tendo sido, posteriormente, adicionado em gelo 12,25µl de uma segunda mistura reacional contendo 4µl de tampão 5x First Strand Buffer (*Invitrogen*) para manter o pH da reacção constante; 4µl de dNTPs (100mM cada, *illustra™ GE Healthcare*); 2µl de ditioneitol (DTT) (0,1M, *Invitrogen*) como estabilizador enzimático; 0,2µl de Single-Stranded DNA, Binding Protein (1-5µg/µl, *Promega*) como estabilizador do cDNA de cadeia simples; 1µl de Superscript® II Reverse Transcriptase (200U/µl, *Invitrogen*), enzima que sintetiza o cDNA a partir de um RNA molde; 0,75µl de RnaseOut™ Recombinant Ribonuclease Inhibitor (40U/µl, *Invitrogen*), para a inactivação de possíveis RNases presentes na reacção e 0,3µl água bidestilada tratada com DEPC de modo a perfazer o volume. Os tubos foram colocados novamente no termociclador tendo sido efectuada uma incubação a 42°C durante 60 minutos seguida de uma etapa de 70°C durante 15 minutos, para inactivação das enzimas. No final da reacção obteve-se um volume final de 20µl de cDNA, o qual foi armazenado a -20°C.

De modo a confirmar o sucesso da reacção de RT-PCR efectuou-se a amplificação por PCR de um fragmento do gene gliceraldeído-3-fosfato desidrogenase (*GAPDH*). Este gene é expresso de forma ubíqua em todas as células do organismo permitindo obter um perfil de bandas homogéneo entre as amostras. A

eficiência do produto de PCR obtido foi controlada em gel de agarose a 2% (p/v) como descrito em 3.2.3.2.

3.2.4.2. Amplificação por PCR do cDNA dos genes *TPP2*, *TEX30* e *SAMSNI*

Após a síntese de cDNA das amostras em estudo e otimização das condições de PCR para os transcritos dos genes *TPP2*, *TEX30* e *SAMSNI* (Figura 9) como descrito em 3.2.2.1, seguiu-se a sua amplificação por PCR (anexo II). A análise mutacional do gene *TPP2* tinha sido realizada em cDNA devido à sua dimensão, tendo este sido dividido em vários fragmentos, os quais foram numerados sequencialmente e não correspondem ao exão amplificado (Pereira 2013). Nas amostras analisadas foram identificados dois transcritos alternativos deste gene, nos fragmentos 5 (transcrito alternativo 13a) e 8 (transcrito alternativo 24a), respectivamente (Figura 9C), sendo necessário verificar se existe alguma relação entre a presença destes e o FCCTX. Para o gene *TEX30* (Figura 9B), foi desenhado um par de *primers* para amplificação da totalidade do transcrito. Enquanto que para o gene *SAMSNI* (Figura 9A), como este possui 3 transcritos codificantes que variam nos exões iniciais do gene, foram desenhados três *primers*: um *forward* para o primeiro exão codificante deste gene (exão 2) e um *reverse* para o exão 9 de modo a amplificar o transcrito 1 e um *primer forward* no exão alternativo dos transcritos 2 e 3 utilizando-se como *primer reverse* o anteriormente mencionado para a amplificação destes.

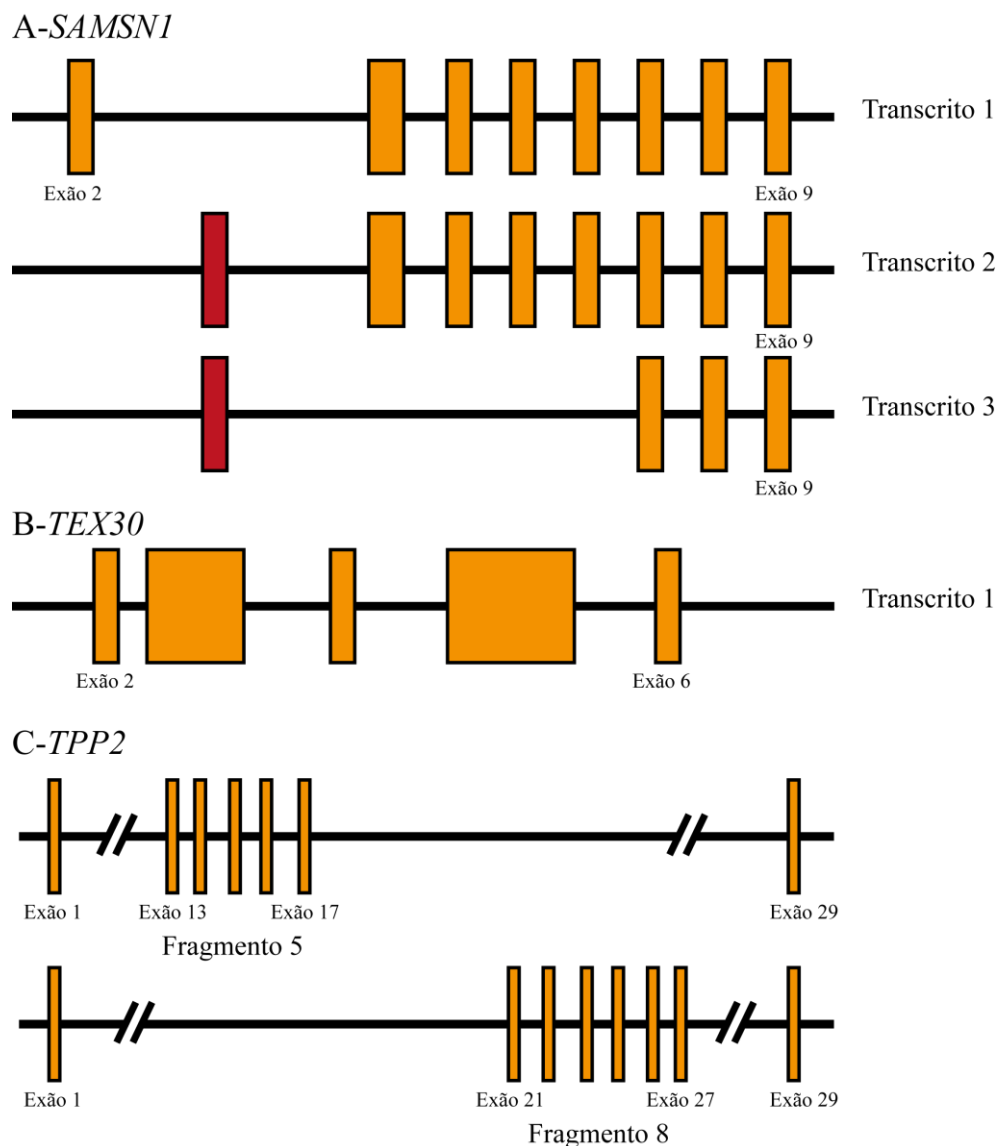


Figura 9. Representação esquemática do cDNA dos genes *SAMSN1* (A), *TEX30* (B), *TPP2* (C). Na secção A, a caixa vermelha representa o exão 1 alternativo existente nos transcritos 2 e 3 do gene *SAMSN1*. Na secção C, encontra-se representado o fragmento 5 do gene *TPP2* contendo os exões 13 a 17 e o fragmento 8 contendo os exões 21 a 27.

O produto amplificado por PCR do fragmento 5 do gene *TPP2*, constituído por 577pb, o fragmento 8 do mesmo gene, constituído por 528pb, e o transcrito 1 do gene *TEX30*, constituído por 930pb foram controlados num gel de agarose 2% enquanto que o transcrito 1 do gene *SAMSN1*, constituído por 1281pb, e o transcrito 2 do mesmo gene, constituído por 1467pb, foram controlados em gel de 1,5%, durante 50 minutos, como descrito em 3.2.2.2. Nos casos em que foram identificadas bandas adicionais ao esperado, foram repetidas as amplificações por PCR. Nas mesmas condições mas para um volume final de 25µl, permitindo obter um maior

rendimento do produto amplificado e procedeu-se ao controlo em gel de agarose como descrito anteriormente. As secções do gel de agarose contendo a banda adicional foi excisada, o cDNA purificado como descrito em 3.2.3.3.2 e procedeu-se à sequenciação automática como descrito em 3.2.3.4.

3.2.4.3. Análise dos transcritos alternativos identificados no contexto do FCCTX

De forma a determinar se o transcrito alternativo identificado no gene *TEX30*, assim como os transcritos alternativos anteriormente identificados no gene *TPP2* (Pereira 2013), se encontram apenas no contexto do FCCTX, estas variantes foram avaliadas em dois grupos distintos de indivíduos. Foram seleccionadas amostras de indivíduos com SL e indivíduos do grupo N. Os cDNA destes indivíduos foram obtidos como descrito em 3.2.4.1 e amplificados por PCR utilizando as condições descritas no anexo II. Os produtos amplificados foram posteriormente controlados em gel de agarose como descrito em 3.2.3.2.

3.2.4.4. Análise da expressão do transcrito alternativo 13a identificado no gene *TPP2* através de PCR semiquantitativo

De forma a determinar se a variabilidade de intensidade das bandas correspondentes ao transcrito alternativo 13a identificado no gene *TPP2*, pode estar relacionada com diferenças de expressão entre os indivíduos, efectuou-se um PCR semi-quantitativo utilizando como gene controlo o *GAPDH*. Para tal foi necessário proceder à optimização do número de ciclos de PCR, utilizando diluições seriadas de cDNA de forma a se obter uma intensidade decrescente e gradual (para metade) à medida que se aumenta a diluição do cDNA para o dobro. Após determinação do número de ciclos óptimo, 26 ciclos para o transcrito do gene *TPP2* e 20 ciclos para o transcrito do gene *GAPDH*, procedeu-se à amplificação por PCR do cDNA dos índices das famílias FCCTX bem como dos cDNA dos indivíduos dos dois grupos descritos em 3.2.4.3. De seguida, controlou-se em gel de agarose a 1,5% (p/v) nas mesmas condições descritas em 3.2.3.2.

3.2.5. *Real-Time* PCR (qPCR)

O *Real-Time* PCR ou PCR quantitativo tem como base o PCR comum, mas para além da análise qualitativa permite também uma análise quantitativa do DNA, apresentando uma maior sensibilidade de amplificação deste. Nesta técnica a amplificação do produto de PCR ocorre em simultâneo com a detecção da quantidade amplificada, não requerendo técnicas adicionais de detecção, como a electroforese em gel de agarose. Esta detecção é obtida com a utilização de uma mistura reaccional que contém um marcador fluorescente. Os marcadores podem ser directos, se se intercalam na dupla cadeia de DNA (SYBR Green I) ou indirectos, como os oligonucleótidos marcados com um fluoróculo que se ligam por complementaridade de bases a uma região específica do DNA de cadeia simples (sondas fluorescentes). A fluorescência obtida através dos marcadores é directamente proporcional à quantidade do produto amplificado por PCR, sendo que ao ciclo em que o sinal de fluorescência ultrapassa o *background*, ou seja, quando se inicia a fase exponencial da amplificação por PCR, é dado o nome de *crossing points* (Cp) (Roche Applied Science 2014). O valor dos Cp pode ser obtido com base em dois métodos, *fit points* ou segunda derivada. O método de *fit points* baseia-se na definição por parte do utilizador de um valor de ordenada (fluorescência), denominado *threshold*, sendo que, o valor de Cp corresponde ao ponto onde a curva de amplificação cruza a recta originada por este (Figura 10).

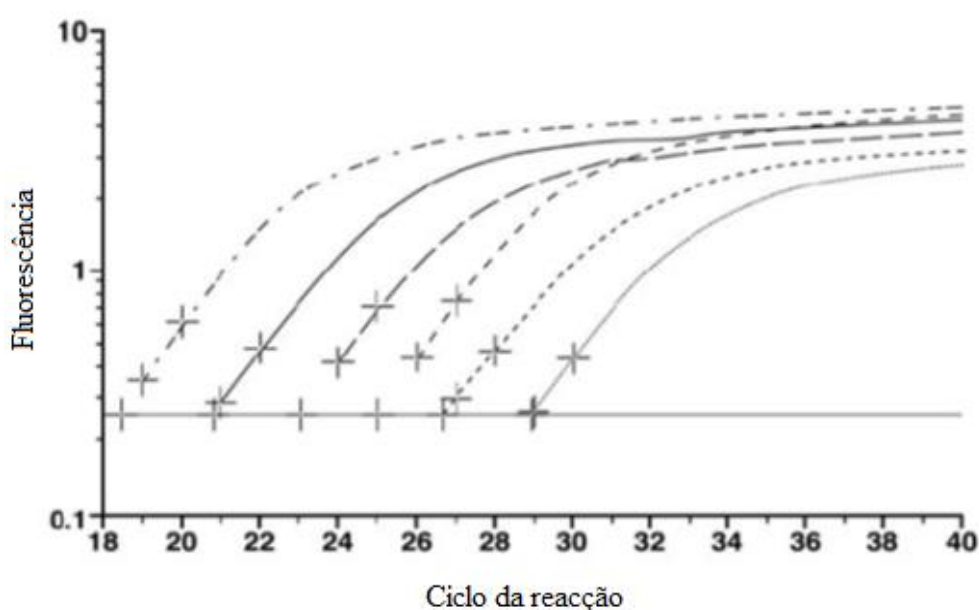


Figura 10. Representação esquemática do método de *fit points* (adaptado de Luu et al. 2005).

O método da segunda derivada (Figura 11) baseia-se no cálculo da segunda derivada da curva de amplificação, sendo obtido um único valor a partir de cada curva. Este método não depende da determinação de um valor de *threshold* pelo utilizador, ou seja, é um método automático e que permite uma melhor reprodutibilidade de resultados.

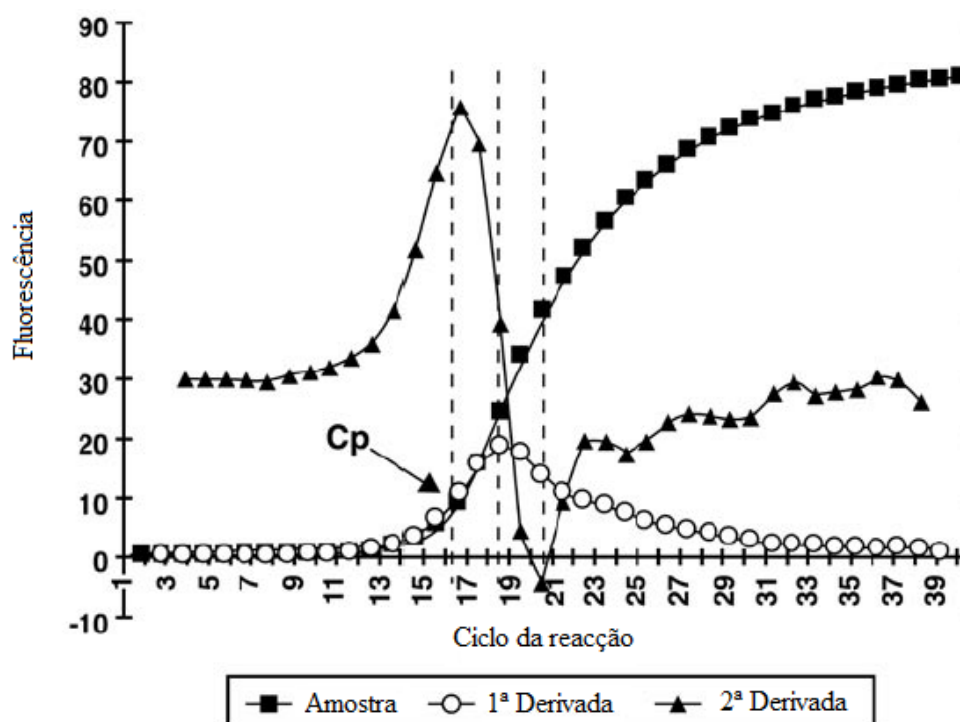


Figura 11. Representação esquemática do método da segunda derivada (adaptado de Luu et al. 2005).

A quantificação das amostras pode ser efectuada por dois métodos, quantificação absoluta ou relativa. Na quantificação absoluta são efectuadas diluições seriadas de uma amostra com concentração conhecida, sendo determinada uma relação linear entre os C_p e o logaritmo da concentração desta amostra, de forma a se obter uma recta de calibração. Partindo desta recta de calibração é possível determinar a concentração de amostras desconhecidas (Roche Applied Science 2014). Na quantificação relativa os valores obtidos (C_p) para as amostras em estudo são inicialmente normalizados contra um gene endógeno, também

denominado de gene *housekeeping*, e posteriormente normalizados contra uma amostra referência (Roche Applied Science 2014).

A especificidade do produto de qPCR é observada através de curvas de *melting*. Após a quantificação, os produtos de PCR são desnaturados com o aumento gradual da temperatura, levando à libertação do marcador fluorescente, o que permite obter um *melting point* específico da sequência amplificada. Para além da especificidade, podemos também usar estas curvas para confirmar o tamanho e o conteúdo em GC do produto de PCR amplificado (Roche Applied Science 2014).

3.2.5.1. Análise da expressão dos genes *TPP2*, *TEX30*, *SAMSN1* e *HSPA13* por Real-Time PCR

O método utilizado para a análise da expressão dos genes em estudo foi um método de quantificação relativa que utiliza o valor da eficiência (E) de cada gene como base da equação:

$$\frac{(E_{\text{Gene alvo}})^{\Delta C_p \text{Amostra}}}{(E_{\text{Gene housekeeping}})^{\Delta C_p \text{Referência}}} \quad \text{Equação 1)}$$

Contudo, é necessário aproximar as eficiências de amplificação do gene alvo e do gene *housekeeping* (Livak & Schmittgen 2001).

3.2.5.1.1. Validação da eficiência de amplificação dos genes em estudo

De forma a aproximar a 100% as eficiências de amplificação dos genes em estudo e dos genes *housekeeping* foi avaliada a variação dos ΔC_p para uma série de diluições seriadas. Para tal foi utilizado uma mistura, em proporções iguais, de 6 amostras de cDNA de indivíduos saudáveis (*pool* de referências normais), sintetizadas como descrito em 3.2.4.1. A *pool* de referências normais foi inicialmente diluída 1:8 em água bidestilada e de seguida foram efectuadas 4 diluições seriadas para metade (1:16 – 1:128). Utilizando cada diluição (1:8 – 1:128) foi efectuado qPCR para cada gene, em duplicado. A média dos C_p dos duplicados foi calculada para cada diluição de cada gene alvo e *housekeeping*. De seguida foi construído um gráfico com o log da concentração de cDNA *versus* ΔC_p (Equação 2).

$$\Delta C_p = C_{p\text{Gene housekeeping}} - C_{p\text{Gene Alvo}} \quad \text{(Equação 2)}$$

A eficiência de cada par de *primers* pode ser calculada individualmente através do declive obtido na equação da recta do gráfico do log da concentração do cDNA *versus* ΔC_p (Rasmussen 2001) e utilizando a equação 3, obtendo-se um valor próximo de 2 = 100% de eficiência.

$$E = 10^{(-1/\text{declive})} \quad (\text{Equação 3})$$

O produto amplificado foi ainda caracterizado recorrendo às curvas de *melting* e sempre que necessário submetido a uma electroforese em gel de agarose a 2% (p/v) como descrito em 3.2.3.2, estes dois métodos permitiram garantir que os produtos amplificados eram específicos e correspondiam ao produto esperado.

As reacção de qPCR foram realizadas no LightCycler® 480 Instrument II (Roche) com o kit *LightCycler®480SYBR Green I Master* em placas de 96 poços utilizando o programa descrito no anexo IV. Cada reacção foi efectuada para um volume final de 15µl, consistindo numa mistura reaccional contendo 0,75µl de cada *primer* à concentração optimizada, 7,5µl de *LightCycler®480SYBR Green I Master* 2x concentrada (Roche) e 4µl de água bidestilada, à qual se adicionou 2µl de cDNA diluído (1:32) em água bidestilada excepto no controlo negativo onde o cDNA foi substituído por água bidestilada.

3.2.5.1.2. Quantificação relativa dos genes em estudo

A quantificação relativa foi efectuada para os transcritos *wild type* dos genes *HSPA13*, *SAMSNI*, *TEX30* e *TPP2*, e ainda para o transcrito alternativo 13a do gene *TPP2*. Para cada transcrito foi desenhado um par de *primers* exónicos, localizados em exões diferentes e sempre que possível na fronteira entre dois exões sequenciais, garantindo especificidade para cDNA. A sequência e a localização dos *primers* utilizados encontra-se no anexo I. Como *housekeeping* foram utilizados os genes, *sucinato desidrogenase subunidade A (SDHA)* e *proteína ribossomal S13 (RPS13)*, os quais foram seleccionados com base no estudo de Duan e colaboradores (Duan et al. 2012).

As reacções de qPCR foram efectuadas em triplicado, em placas de 96 poços, como descrito no ponto anterior, para os genes em estudo em 6 indivíduos da família L55 e 1 indivíduo da família L7, ambas famílias FCCTX das quais existia material biológico disponível. Para além destes indivíduos incluíram-se ainda indivíduos do

grupo N e indivíduos com SL para o estudo dos transcritos dos genes *HSPA13*, *SAMSNI* e *TPP2*. Em cada placa foram efectuadas, ainda, reacções em duplicado para cada gene utilizando o cDNA da *pool* de referências normais, para posterior normalização dos resultados. O cDNA dos indivíduos em estudo foi sintetizado como descrito em 3.2.4.1.

Os valores de Cp obtidos através do *software* Light Cycler 480® v1.5.0.39 (Roche) foram importados para uma folha de cálculo de Excel e a normalização foi realizada de acordo com as equações 1 e 2, utilizando a média dos Cp calculada a partir dos triplicados de cada reacção.

A análise estatística foi realizada no *software* GraphPad Prism version 5.03, utilizando a média e o desvio padrão dos triplicados de cada experiência e recorrendo a um teste estatístico t.

3.2.5.2. Análise de *copy number* nos genes *TPP2*, *SAMSNI* e *HSPA13* nas famílias FCCTX por qPCR

De modo a se proceder à análise de *copy number* dos genes em estudo, foram aproximadas as eficiências de amplificação utilizando o mesmo método de diluições seriadas aplicado no ponto 3.2.5.1.1 mas utilizando uma mistura de amostras de DNA (*pool* de referências normais). Neste caso, o DNA *stock* foi quantificado em Nanodrop 2000 (*Thermo Scientific*) e efectuou-se uma diluição seriada para 80ng/μl, a partir desta diluição efectuaram-se 5 diluições seriadas para metade (1:2 – 1:32).

Após a optimização das eficiências de amplificação dos *primers*, procedeu-se à quantificação relativa dos genes *TPP2*, *SAMSNI* e *HSPA13*. Para o gene *TPP2* tinham sido desenhados, previamente a este estudo, dois pares de *primers* descritos no anexo I. O primeiro par (*TPP2* intrão 6) para a fronteira entre a região intrónica 6/exão 7 e exão 7 e o segundo par (*TPP2* intrão 21 - D13S1256) para a região intrónica 21, este coincidente com a região amplificada pelo marcador de microssatélites D13S1256. Relativamente aos outros dois genes, *SAMSNI* e *HSPA13*, foi desenhado um par de *primers*, para cada, sendo que a sequência e a localização destes se encontra descrito no anexo I. Como gene *housekeeping* foi utilizada uma região do cromossoma 12 (CHR12) compreendida entre 6585217pb e 6585317pb, utilizada numa análise de *copy number* por Duan e colaboradores (Duan et al. 2012).

As reacções de qPCR foram efectuadas em triplicado para os genes em estudo, em placas de 96 poços, como descrito em 3.2.5.1.1, utilizando 2µl de DNA diluído (1:4 ≈ 20ng/µl) em água bidestilada, para os mesmos 6 indivíduos da família L55 e 1 indivíduo da família L7 para os quais tinha sido efectuada a análise de expressão nestes genes. Em cada placa foi ainda incluída uma reacção em duplicado para cada gene, utilizando a *pool* de DNA de referências normais, para posterior normalização dos resultados.

A análise estatística foi realizada como descrito em 3.2.5.1.2.

3.2.5.3. Análise de *copy number* na região 13q32-33 em tumores de indivíduos FCCTX por qPCR

Para a análise da variabilidade de *copy number* da região 13q32-33 seleccionaram-se dois genes de interesse nesta região, nomeadamente, os genes *TPP2* e *DOCK9*. Para o gene *TPP2* foram utilizados os mesmos *primers* mencionados em 3.2.5.2, enquanto que para o gene *DOCK9* foi utilizado um par de *primers*, desenhados previamente a este estudo, para a região intrónica 2, este gene encontra-se próximo da região amplificada pelo marcador de microssatélite D13S1298. Como mencionado anteriormente, foi necessário aproximar as eficiências de amplificação utilizando o mesmo método de diluições seriadas aplicado no ponto 3.2.5.1.1 e 3.2.5.2, tendo sido utilizada uma amostra de DNA de mucosa normal incluída em parafina.

Após a optimização das eficiências de amplificação dos *primers*, procedeu-se à quantificação relativa dos genes *TPP2* e *DOCK9*, utilizando-se o mesmo gene *housekeeping* utilizado para a análise de *copy number* descrita em 3.2.5.2.

As reacções de qPCR foram realizadas em triplicados, em placas de 96 poços, como descrito em 3.2.5.1.1, utilizando 2µl de DNA de FFPE diluído (1:4 ≈ 20ng/µl) em água bidestilada, para os genes em estudo em 12 tumores FCCTX dos quais dois pertenciam ao mesmo indivíduo, e sempre que possível foi também analisada a mucosa normal respectiva (10/12). Em cada placa foi ainda incluída uma reacção em duplicado para cada um dos genes, utilizando duas amostras de DNA de mucosa normal (referencias normais), para posterior normalização.

A análise estatística foi realizada como descrito em 3.2.5.1.2.

3.2.6. Análise das mutações c.292 G>A, c.868 C>T e c.511 G>T no gene *SLC10A2* no contexto do FCCTX

Na análise mutacional para o gene *SLC10A2*, realizada anteriormente a este estudo, foram detectadas três mutações pontuais: no exão 1 c.292 G>A, no exão 3 c.511 C>T e no exão 5 c.868 C>T (Pereira 2013). De forma a perceber se alguma destas mutações pode estar relacionada com a carcinogénese do FCCTX efectuou-se a pesquisa das mutações, c. 292 G>A e c.868 C>T, em 48 indivíduos saudáveis de modo a avaliar a frequência destas mutações na população em geral. Em relação à mutação encontrada no exão 3, esta tinha sido detectada no indivíduo índice da família L84 (da qual se teve acesso ao DNA de vários familiares (afectados, em risco e saudáveis) os quais já tinham sido analisados num trabalho anterior (Francisco et al. 2011). Assim para esta família efectuou-se uma análise de *linkage*, de forma a verificar se a mutação segrega com a doença. As amostras de DNA foram obtidas através de sangue periférico como descrito em 3.2.1.1, as condições de amplificação por PCR para sequenciação automática dos três exões encontram-se descritas no anexo II.

4. Resultados

4.1. Envolvimento de genes candidatos na susceptibilidade para o FCCTX

Os estudos realizados por Nieminen e colaboradores (Nieminen et al. 2011) identificaram a presença de mutações germinais no *BMPRIA* como uma possível causa molecular para o cancro do cólon e recto familiar em algumas famílias com tumores MSS, o que tornou este gene num possível candidato para a susceptibilidade para o FCCTX. Por outro lado, foram identificadas em estudos prévios, realizados pelo grupo de Gastrenterologia da UIPM, 2 regiões de susceptibilidade para o FCCTX (13q32-33 e 21q11). Nestas 2 regiões de susceptibilidade foram seleccionados genes candidatos tendo em conta os critérios referidos anteriormente em 1.4.2. Na região 21q11, após a exclusão de mutações no gene *NRIP1*, os genes *HSPA13* e *SAMSNI*, pela sua função já explicada em 1.4.3.1 e 1.4.3.2, respectivamente, são também fortes genes candidatos, tendo por isso sido analisado o seu envolvimento através da pesquisa de mutações pontuais, grandes deleções/inserções e alterações de expressão, ao nível germinal.

No caso da região 13q32-33, a análise do envolvimento na susceptibilidade para o FCCTX de alguns genes candidatos em estudos prévios não excluiu por completo o gene *TEX30* ao nível da alteração de expressão ou de grandes deleções/inserções e levantou-se a hipótese de algumas alterações nos genes *TPP2* (expressão de um transcrito alternativo) e *SLC10A2* (mutações p. V98I, p. P290S e p.A171S) poderem estar associadas a este síndrome. O presente estudo pretende contribuir para esclarecer o envolvimento destes genes/alterações genéticas.

4.1.1. *BMPRIA*

Os resultados da análise mutacional dos 11 exões e respectivas regiões flanqueadoras do gene, por indivíduo e por exão, encontram-se no anexo VII. Foram identificadas 2 mutações exónicas, uma *missense* no exão 3 (p.Pro2Thr) com uma frequência alélica de 23% (Tabela 5) e outra silenciosa no exão 10 (p.Asp380Tyr) com uma frequência de 3%, 2 mutações intrónicas, uma no intrão 6 (c.431-26 T>A) com uma frequência de 3% e outra no intrão 11 (c. 448-11 T>C) com uma frequência de 23% e 4 na região 3'UTR (c.1599+847 G>A, c.1599+764 T>C, c.1599+886 C>T

e c.1599+1497 T>C) com frequências de 3%, 9%, 21% e 12%, respectivamente . Nesta região foi ainda encontrada uma mutação do tipo *frameshift* (c. 1577_1580delTTTA) não descrita na base de dados *Gene* do NCBI (<http://www.ncbi.nlm.nih.gov/gene>) e com uma frequência de 18% nos indivíduos FCCTX analisados.

Tabela 5. Caracterização das mutações identificadas no gene *BMPRIA*.

Localização	Mutação germinal	Frequência alélica				Variant Effect Predictor		Mutation Taster	Referência bibliográfica
		FCCTX		População europeia		SIFT	Polyphen		
Exão 3	c. 4 C>A p.Pro2Thr (rs11528010)	A	24%	A	25%	Tolerado (1)	Benigno (0)	Polimorfismo	Kaczmarczyk et al., 2013 (PMID: 22886282); Böttcher Y. et al., 2009 (PMID: 19502417).
Intrão 6	c.431-26 T>A (rs7920259)	A	3%	A	1%	Variante intrónica		Polimorfismo	—
Exão 10	c. 1140 C>T p.Asp380Tyr (rs35572415)	T	3%	T	7%	Variante sinónima		Patogénica * ¹	—
Intrão 11	c.448-11 T>C (rs7074064)	C	24%	C	28%	Variante intrónica		Polimorfismo	—
3'UTR	c.1599+847 G>A	A	3%	—		Variante na região 3'UTR		Patogénica * ²	—
	c.1599+764 T>C (rs150762162)	C	9%	C	11%	Variante na região 3'UTR		Polimorfismo	—
	c.1599+886 C>T (rs149499991)	T	21%	T	10%	Variante na região 3'UTR		Patogénica * ³	—
Região intergénica	c.1599+1497 T>C (rs17335915)	C	12%	C	27%	Variante intergénica		—	—
	c.*1577_*1580delTTTA (rs150276085)	del	18%	del	14%	Variante intergénica		—	—

*¹Criação de um novo local de *splicing* 2pb após a mutação (wt: 0,29 / mut: 0,41); *² Criação de um novo local de *splicing* 4pb antes da mutação (wt:0,66 / mut: 0,98);

*³Criação de um novo local de *splicing* 9pb antes da mutação (wt: 0,93 / mut: 0,96).

A análise *in silico*, utilizando os *softwares* *Variant effect predictor* (McLaren et al. 2010) e *Mutation taster* (Schwarz et al. 2010), revelou 3 mutações com carácter possivelmente patogénico. As mutações A380T, c.1599+847 G>A e c.1599+886 C>T podem levar à criação de novos locais de *splicing* na proximidade da respectiva mutação.

Destas, a mutação c.1599+886 C>T apresentou uma elevada frequência (21%) quando comparada com a frequência na população europeia (10%), no entanto na população ibérica a frequência alélica é próxima da constatada (18%) o que suporta tratar-se de uma variante polimórfica.

Apesar da análise *in silico* não prever um carácter patogénico para a mutação c.1599+1497, esta revelou uma baixa frequência (12%) quando comparada com a população europeia (27%).

4.1.2. *HSPA13*

Os resultados da análise mutacional dos 5 exões e respectivas regiões flanqueadoras do gene, por indivíduo e por exão, encontram-se no anexo VII. Esta análise, revelou a presença de apenas 2 mutações em regiões intrónicas, uma no intrão 3 (c.581-46 G>T) e outra no intrão 4 (c.749-56 G>A) (Tabela 6). A primeira encontra-se descrita como rs2822642 e possui uma frequência alélica de 59% versus 77% na população europeia, não se verificando diferenças relevantes. A mutação c.749-56 G>A, não se encontra descrita na base de dados *Gene* do NCBI e a frequência nos indivíduos FCCTX analisados foi de 3%.

Tabela 6. Caracterização das mutações identificadas no gene *HSPA13*.

Localização	Mutação germinal	Frequência alélica				Variant Effect Predictor		Mutation Taster
		Indivíduos FCCTX		População europeia		SIFT	Polyphen	
Intrão 3	c.581-46 G>T (rs2822642)	T	59%	T	77%	Variante intrónica		Polimorfismo
Intrão 4	c.749-56 G>A	A	3%	—		Variante intrónica		Polimorfismo

A análise *in silico* revelou que nenhuma das mutações identificadas parece apresentar um carácter patogénico. Adicionalmente a elevada frequência da mutação c. 581-46 G>T, quer na população em estudo quer na população europeia suporta o seu carácter não patogénico.

De modo a excluir o envolvimento deste gene na susceptibilidade para o FCCTX, é necessário avaliar a presença de deleções de 1 ou mais exões, alteração que não é detectada por sequenciação automática, no entanto não foi possível efectuar esta análise ao nível do RNA.

Foi também avaliada a existência de alterações na expressão do gene que pudessem ser responsáveis pelo fenótipo. Verificou-se que 2/6 dos indivíduos analisados, pertencentes à mesma família (L55) (Figura 12), apresentavam um aumento significativo de expressão do transcrito deste gene em relação ao gene *housekeeping* (L1544 e L447), os restantes indivíduos analisados não apresentaram alterações de expressão deste gene.

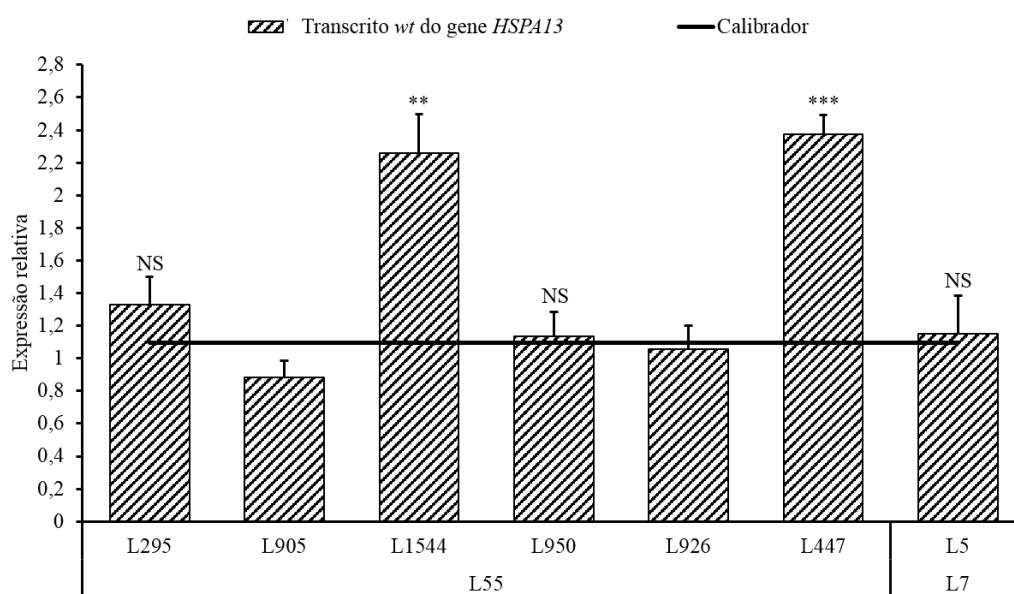


Figura 12. Resultados da expressão relativa do transcrito *wt* do gene *HSPA13* para os indivíduos de 2 famílias FCCTX (L55 e L7). Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). *** $p < 0,001$; ** $p < 0,01$; NS – não significativo.

Tendo em conta o aumento de expressão do gene em dois indivíduos de uma das famílias FCCTX, foi avaliada a expressão deste transcrito em indivíduos saudáveis (grupo N) e também em indivíduos com SL, de modo a aumentar o número de indivíduos analisados com uma expressão esperada normal, uma vez que nestes a causa molecular da doença já se encontra identificada. Não se verificaram alterações relevantes ao nível da expressão deste transcrito em nenhum destes grupos de indivíduos (Figura 13).

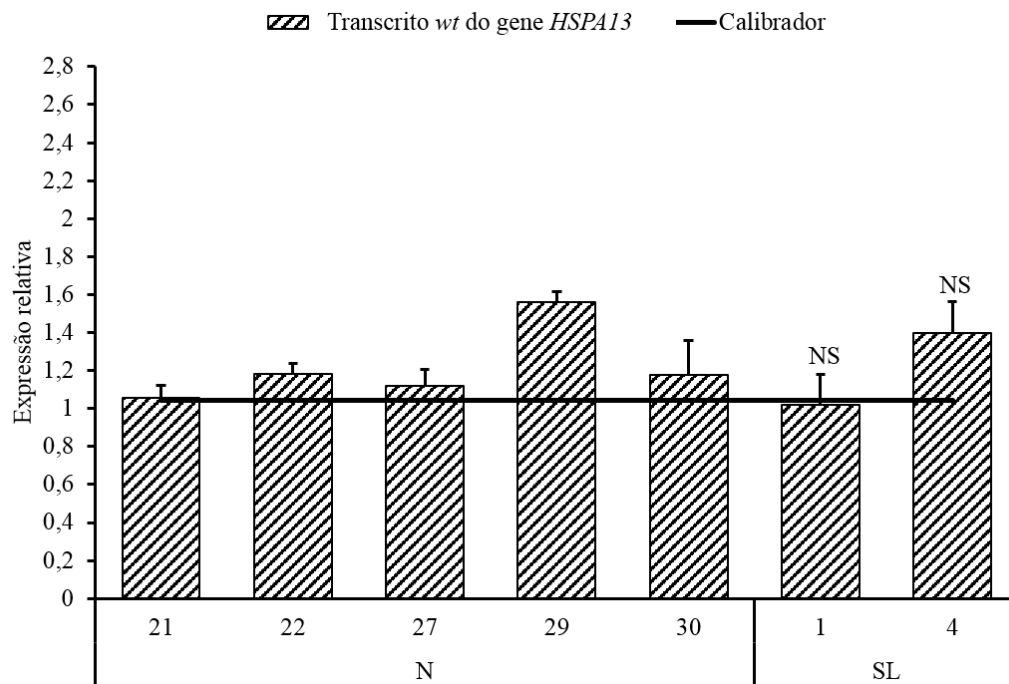


Figura 13. Resultados da expressão relativa do transcrito wt do gene *HSPA13* em indivíduos saudáveis (N) e em indivíduos com SL. Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). NS – não significativo.

De modo a excluir a existência de uma deleção total do gene, foi realizada a análise de *copy number* utilizando *primers* específicos descritos no anexo I. Os resultados da análise de *copy number*, não revelaram deleção deste gene em nenhum dos indivíduos das famílias analisadas (L55 e L7), sendo os valores semelhantes à média apresentada no grupo controlo (Calibrador) (Figura 18).

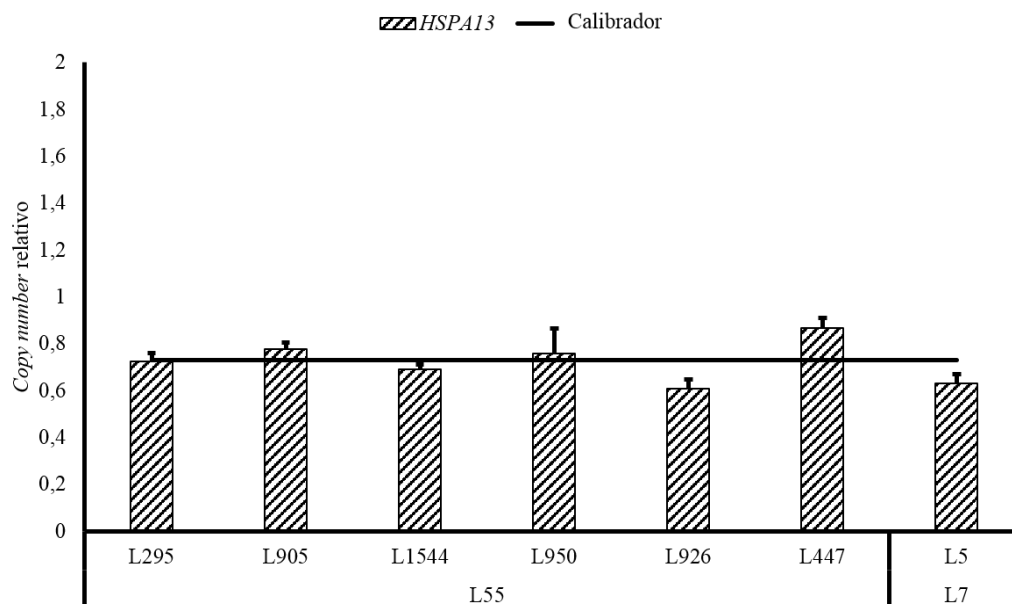


Figura 14. Resultados da análise de *copy number* do gene *HSPA13* para os indivíduos de 2 famílias FCCTX (L55 e L7). Calibrador – Média do *copy number* relativo do grupo controlo.

4.1.3. *SAMSNI*

Os resultados da análise mutacional dos 8 exões e respectivas regiões flanqueadoras do gene *SAMSNI*, realizada para os vários indivíduos das famílias FCCTX encontram-se em detalhe no anexo VII. Estes resultados revelaram a presença de 3 mutações *missense* no exão 2 (p.Asn20Asp, p.Pro42Leu e p.His64Tyr) e 3 mutações intrónicas, 2 no intrão 2 (c.262-52 A>G, c.262-25 T>A) e 1 no intrão 8 (c.920-124 C>T) (Tabela 7). Todas as mutações apresentaram uma frequência relativamente elevada e encontram-se descritas na base de dados *Gene* do NCBI. À excepção das mutações, H64Y e c.262-52 A>G, as restantes apresentam frequências alélicas nos indivíduos FCCTX semelhantes às da população europeia. A mutação H64Y possui uma frequência alélica de 38% *versus* 57% na população europeia,

enquanto que a segunda mutação, c.262-52 A>G, possui uma frequência de 32% *versus* 54% na população europeia.

Tabela 7. Caracterização das mutações identificadas no gene *SAMSN1*.

Localização	Mutação germinal	Frequência da mutação				Variant Effect Predictor		Mutation Taster
		FCCTX		População europeia		SIFT	Polyphen	
Exão 2	c. 58 A>G p.Asn20Asp (rs2822787)	G	91%	G	81%	Tolerado (0,69)	Benigno (0)	Polimorfismo
	c. 71 C>T p.Pro42Leu (rs2822786)	A	35%	A	34%	Tolerado (0,27)	Benigno (0,01)	Polimorfismo
	c. 190 C>T p.His64Tyr (rs2822785)	A	38%	A	57%	Tolerado (1)	Benigno (0)	Polimorfismo
Intrão 2	c.262-52 A>G (rs760345)	G	32%	G	54%	Variante intrónica		Polimorfismo
	c.262-25 T>A (rs760344)	A	21%	A	15%	Variante intrónica		Polimorfismo
Intrão 8	c.920-124 C>T (rs3746826)	T	9%	T	7%	Variante intrónica		Polimorfismo

A análise *in silico* revelou que nenhuma das mutações identificadas parece apresentar carácter patogénico, o que está de acordo com a elevada frequência descrita.

A análise de grandes deleções revelou que para além do transcrito *wt* detectado em todas as amostras, foi possível visualizar um transcrito de menor peso molecular o qual parecia apresentar intensidade semelhante em todas as amostras (Figura 15), no entanto ainda não foi possível sequenciar o mesmo de forma a identificar a que transcrito alternativo corresponde. Constataram-se também diferenças de intensidade no transcrito *wt* entre as várias amostras analisadas, quando comparado com o gene *housekeeping* utilizado, *GAPDH*. Assim, procedeu-se à realização de uma análise de expressão deste transcrito por qPCR de modo a esclarecer estas diferenças.

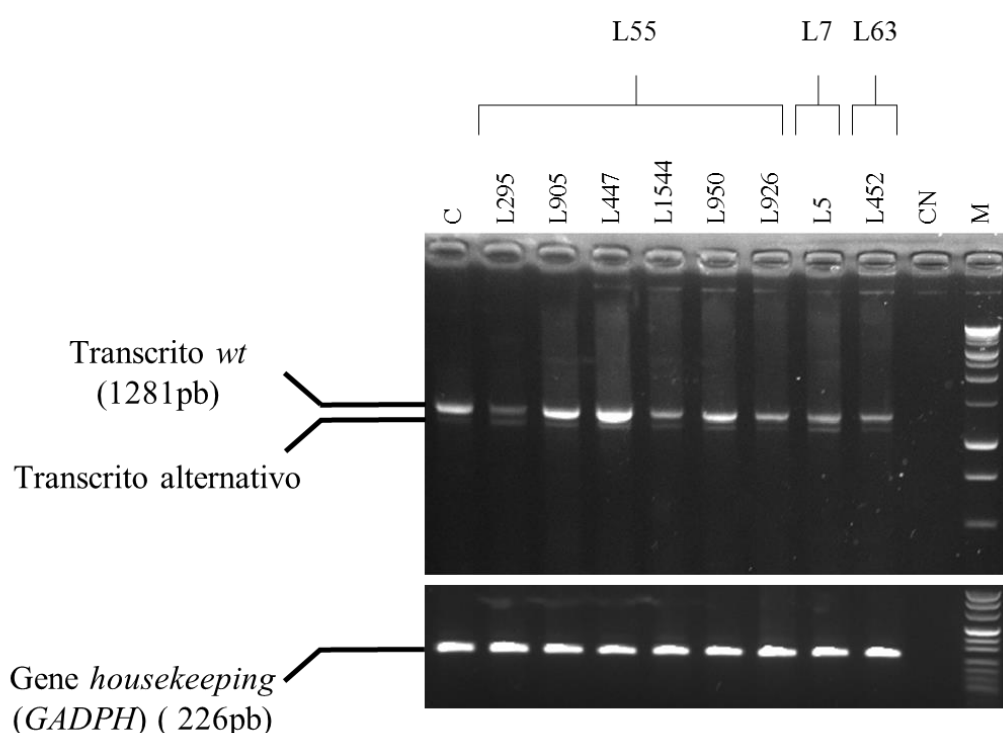


Figura 15. Electroforese em gel de agarose 1,5% (p/v) dos produtos amplificados por PCR para o gene *SAMSNI* e 2% (p/v) para o gene *GAPDH*. C – Controlo normal; CN – Controlo negativo; M – Marcador molecular *1kb DNA Ladder* (Promega) (Topo) e *GeneRuler 50bp DNA Ladder* (Thermo Scientific) (fundo).

Os resultados da análise de expressão do transcrito *wt* deste gene, revelaram um aumento significativo de expressão deste gene em relação ao gene *housekeeping* em 2/7 dos indivíduos analisados (L1544 e L447), ambos da família L55 (Figura 16). Esta análise revelou ainda uma diminuição significativa de expressão deste gene em 1/7 dos indivíduos da família L55 (L295).

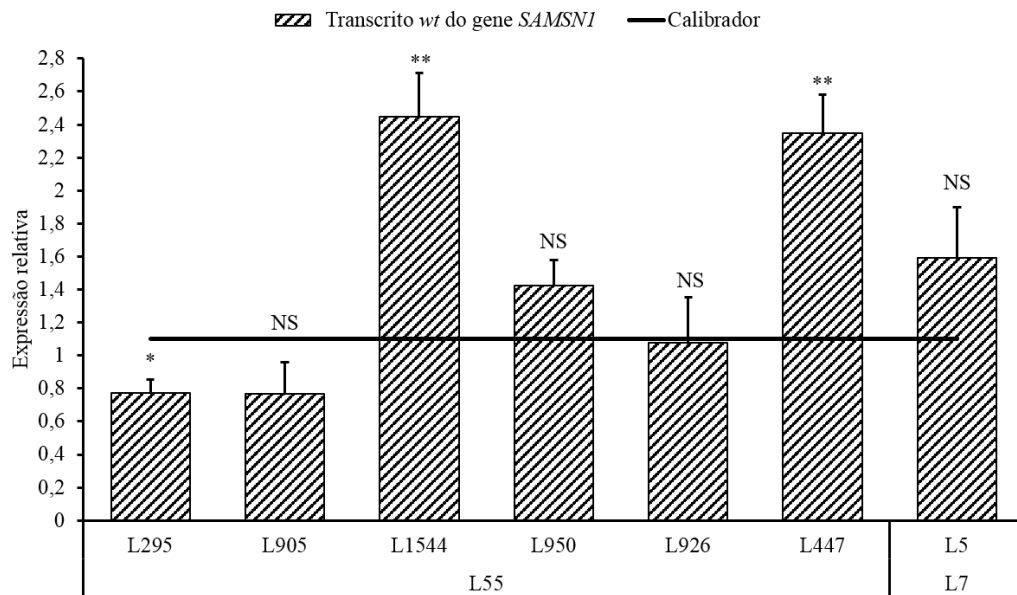


Figura 16. Resultados da expressão relativa do transcrito *wt* do gene *SAMSNI* para os indivíduos de 2 famílias FCCTX (L55 e L7). Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). ** $p < 0,01$; * $p < 0,05$; NS – não significativo.

A análise de expressão deste transcrito em indivíduos saudáveis (N) e indivíduos com SL revelou um aumento significativo de expressão em um dos dois indivíduos de síndrome de Lynch (1) (Figura 17). Os restantes indivíduos apresentaram valores semelhantes à média apresentada no grupo controlo (Calibrador).

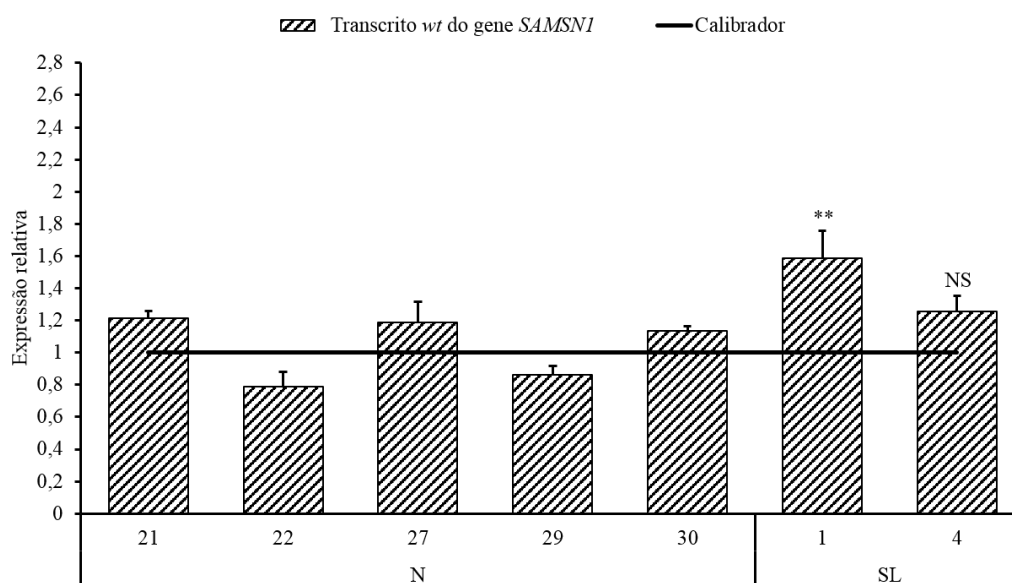


Figura 17. Resultados da expressão relativa do transcrito *wt* do gene *SAMSNI* em indivíduos saudáveis (N) e em indivíduos com SL. Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). ** $p < 0,001$; NS – não significativa.

De modo a excluir a existência de uma deleção total deste gene, foi realizada a análise de *copy number* utilizando *primers* específicos descritos no anexo I. Os resultados desta análise revelaram uma diminuição do número de cópias deste gene em dois dos indivíduos FCCTX (L926 e L5), a qual foi significativa apenas no primeiro caso (Figura 18). Relativamente ao L5 não foi possível repetir a amostra de forma a obter valores significativos para esta alteração. Verificou-se ainda que um dos indivíduos FCCTX (L447), da mesma família que o L926 apresentou um aumento do número de cópias deste gene, apesar de não se ter conseguido obter valores significativos para esta alteração.

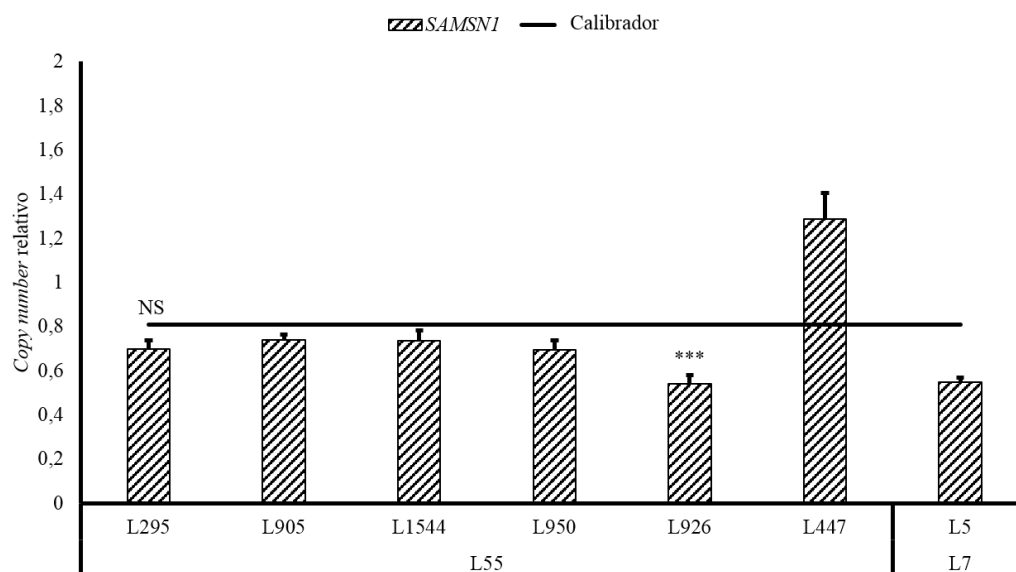


Figura 18. Resultados da análise de *copy number* do gene *SAMSNI* para os indivíduos de 2 famílias FCCTX (L55 e L7). Calibrador – Média do *copy number* relativo do grupo controle. *** $p < 0,0001$; NS – não significativo.

4.2. Análise do envolvimento de genes candidatos localizados na região de susceptibilidade 13q32-33

4.2.1. *TPP2*

Num estudo prévio, a análise mutacional do gene *TPP2* revelou a presença de dois transcritos alternativos deste gene, um que apresenta uma inserção de 72 nucleótidos entre o exão 13 e 14 (transcrito alternativo 13a), e outro contendo uma inserção de 39 nucleótidos entre os exões 23 e 24 (transcrito alternativo 24a). O transcrito alternativo 24a encontra-se descrito como sendo o transcrito 2 do gene *TPP2*, segundo a base de dados *Ensembl*, o qual é constituído por 4660pb e codifica para uma proteína com 1262 aminoácidos, mais 13 do que o transcrito *wt*. No mesmo estudo foi avaliada a expressão dos dois transcritos em indivíduos de famílias FCCTX, indivíduos com SL e indivíduos saudáveis (N) tendo sido observada a redução ou ausência de expressão do transcrito *wt* em alguns indivíduos FCCTX, assim como em alguns indivíduos com SL (Pereira 2013). No presente estudo foi alargada a análise de expressão destes transcritos a um maior número de amostras de cada um dos grupos.

De acordo com os resultados obtidos na electroforese em gel de agarose 1,2% (p/v), verificaram-se diferenças de expressão de ambos os transcritos, *wt*, 13a e 24a, nos três grupos analisados, famílias FCCTX, indivíduos com SL e indivíduos do grupo N.

Relativamente ao transcrito 24a, um estudo anterior tinha revelado, aparentemente, diferenças de expressão entre os dois transcritos (Pereira 2013). No presente estudo a análise de um maior número de amostras revelou que a maioria apresentava uma maior expressão do transcrito *wt* em relação ao transcrito 24a, quer nos indivíduos de famílias de FCCTX, indivíduos com SL ou indivíduos saudáveis (Figura 19).

Em relação ao transcrito 13a, verificou-se que nas famílias FCCTX 4/8 (50%) dos indivíduos apresentaram um aumento de expressão do transcrito alternativo 13a (L295, L447, L1544 e L950, todos da mesma família) e nenhum dos indivíduos analisados apresentou um aumento de expressão do transcrito *wt*. Salienta-se ainda que 3/4 (75%) dos indivíduos com um aumento de expressão do transcrito 13a apresentam uma redução quase total de expressão do transcrito *wt* (L447, L1544 e L950).

A análise do grupo de indivíduos saudáveis permitiu identificar um aumento da expressão do transcrito 13a em 4/15 (27%) dos indivíduos (31-34), com redução quase total de expressão do transcrito *wt* em três destes. (Figura 19).

A análise adicional em indivíduos com SL revelou um aumento de expressão do transcrito 13a em 7/19(37%) (2, 4, 6, 10, 14, 17 e 19) enquanto que 6/19 (32%) apresentaram um aumento de expressão do transcrito *wt*. Salienta-se ainda que 5/7 (71%) indivíduos com aumento de expressão do transcrito 13a apresentaram uma redução quase total de expressão do transcrito *wt* (2, 4, 6, 10 e 19).

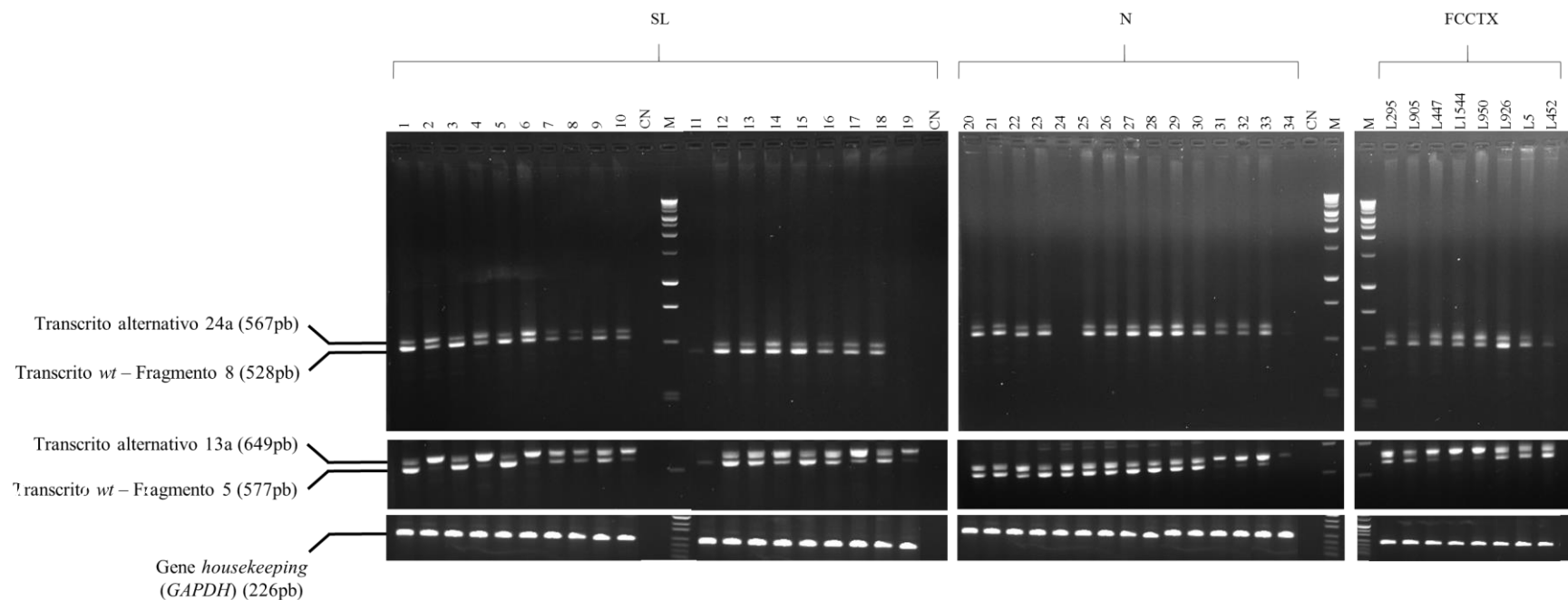


Figura 19. Electroforese em gel de agarose 1,2% (p/v) dos produtos amplificados por PCR para o transcrito alternativo 24a (topo), transcrito alternativo 13a (meio) do gene *TPP2* e 2% (p/v) para o gene *housekeeping GAPDH* (fundo) para os indivíduos de famílias FCCTX, indivíduos com SL e indivíduos saudáveis (N). Devido à sua dimensão o gene *TPP2* foi dividido em 9 fragmentos para amplificação do cDNA: Fragmento 5 – Exão 13 a 17; Fragmento 8 – Exão 21 a 27. CN – Controlo negativo; M – Marcador de peso molecular *1kb DNA Ladder* (Promega) (topo e meio) e *GeneRuler 50bp DNA Ladder* (Thermo Scientific) (fundo).

Tendo-se verificado diferenças de expressão no transcrito alternativo 13a e uma vez que este não se encontra descrito, procedeu-se à realização de um PCR semi-quantitativo, com o intuito de confirmar que as diferenças de intensidade dos fragmentos obtidos por PCR correspondiam de facto a uma expressão diferencial dos transcritos.

Os resultados obtidos confirmaram que 2 dos indivíduos FCCTX, ambos da mesma família, apresentaram um aumento de expressão do transcrito 13a e quase ausência de expressão do transcrito *wt* (L447 e L950). O mesmo padrão foi observado em 3/19 (16%) dos indivíduos com SL e em 2/15 (13%) dos indivíduos saudáveis (2, 10, 17, 32 e 33) (Figura 20). Esta quase ausência de expressão do transcrito *wt*, com aumento da expressão do transcrito 13a, aliada ao facto deste não estar descrito levou a considerar importante a quantificação da expressão por PCR quantitativo.

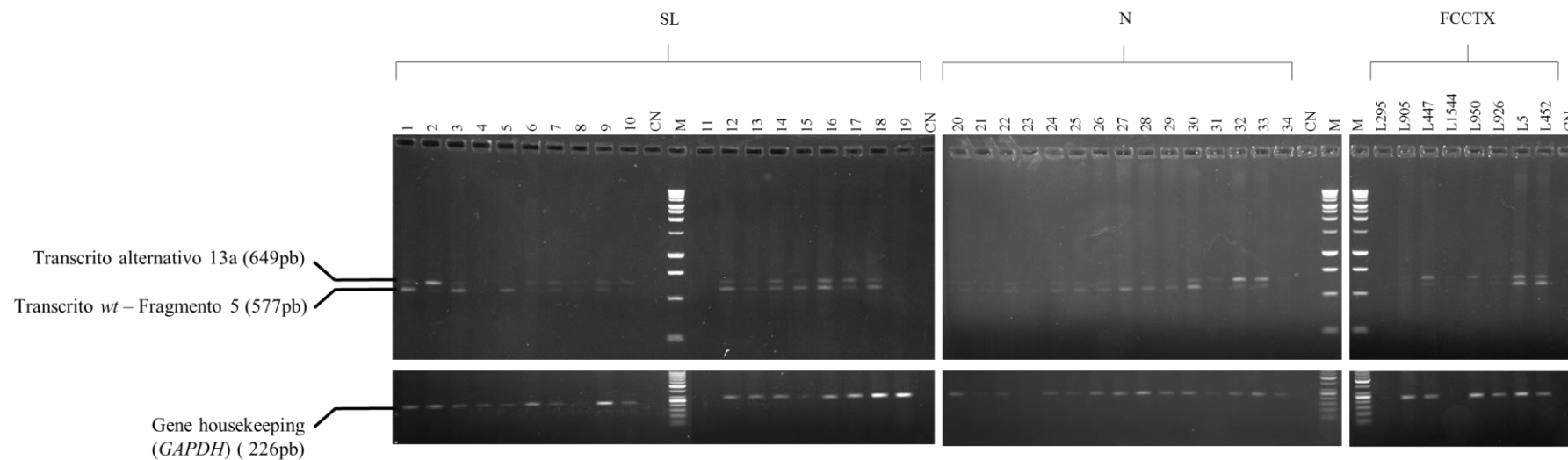


Figura 20. Electroforese em gel de agarose 1,2% (p/v) dos produtos amplificados por PCR semi-quantitativo para o transcrito alternativo 13a (topo) do gene *TPP2* e 2% (p/v) para o gene *housekeeping GAPDH* (fundo)) para os indivíduos de famílias FCCTX, indivíduos com SL e indivíduos saudáveis (N). Devido à sua dimensão o gene *TPP2* foi dividido em 9 fragmentos para amplificação do cDNA: Fragmento 5 – Exão 13 a 17. CN – Controle negativo; M – Marcador de peso molecular *1kb DNA Ladder (Promega)* (topo e meio) e *GeneRuler 50bp DNA Ladder (Thermo Scientific)* (fundo).

Os resultados desta análise confirmaram que 3/7 (43%) dos indivíduos FCCTX apresentaram um aumento de expressão do transcrito 13a e redução acentuada do transcrito *wt* (L1544, L950 e L447, todos da mesma família). Estes resultados revelaram ainda sobreexpressão do transcrito 13a, independentemente da redução do *wt* em 5/7 (71%) indivíduos FCCTX, sendo todos da família L55, no entanto os resultados do indivíduo L295 não foram significativos devido à variabilidade dos valores obtidos (Figura 21).

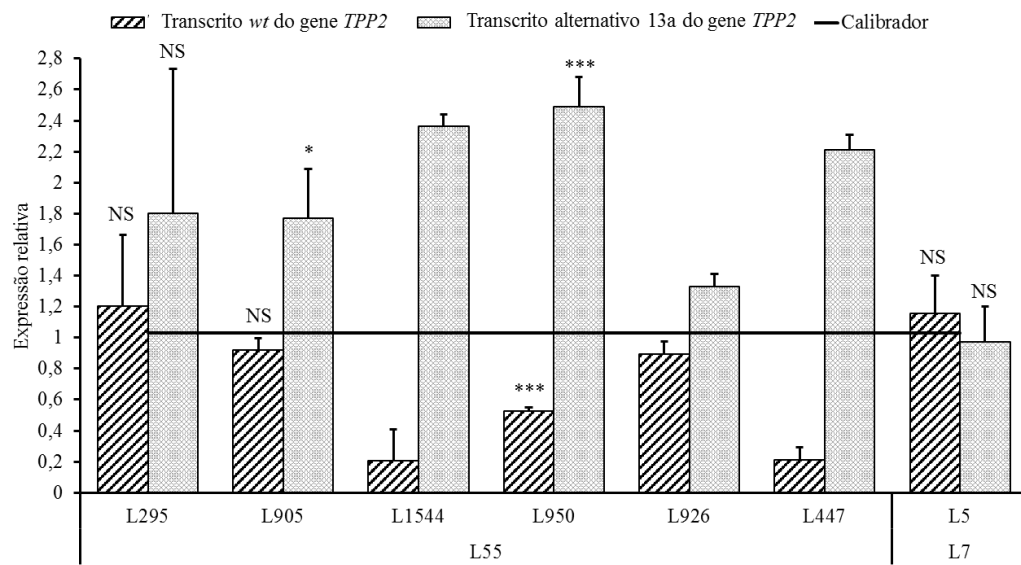


Figura 21. Resultados da expressão relativa do transcrito *wt* e do transcrito 13a do gene *TPP2* para os indivíduos de 2 famílias FCCTX (L55 e L7). Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). *** $p < 0,001$; * $p < 0,05$; NS – não significativo.

Nos indivíduos saudáveis verificou-se a sobreexpressão do transcrito 13a e redução do transcrito *wt* em apenas 3/17 (18%) indivíduos. Constatou-se ainda que 6/17 (35%) dos indivíduos apresentaram um aumento de expressão do transcrito 13a independentemente de apresentarem alteração do transcrito *wt*. Identificaram-se ainda 5 indivíduos com expressão aumentada do transcrito *wt* (Figura 22).

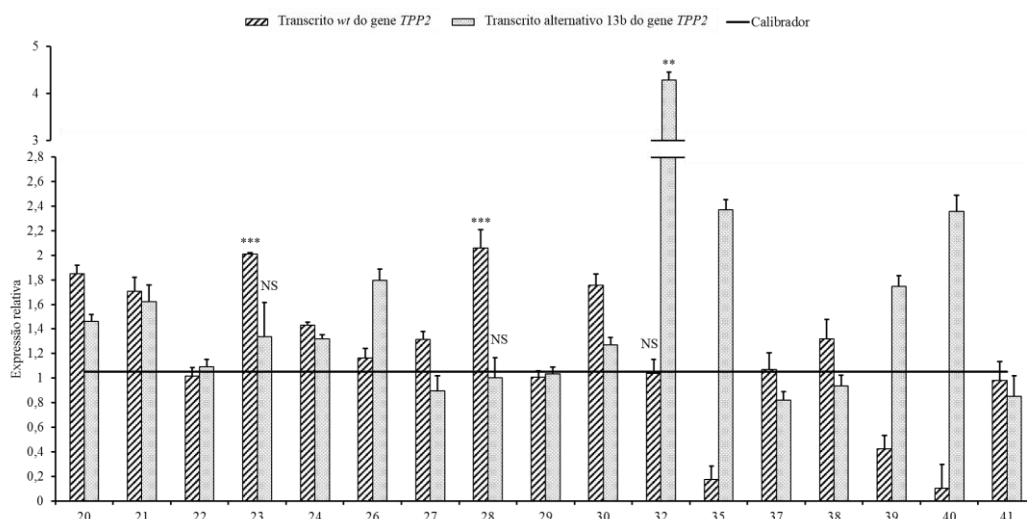


Figura 22. Resultados da expressão relativa do transcrito *wt* e do transcrito 13a do gene *TPP2* em indivíduos saudáveis (N). Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). *** $p<0,001$; ** $p<0,01$; NS – não significativo.

De forma a avaliar a existência desta alteração de expressão noutro grupo de indivíduos, não relacionados com o FCCTX, procedeu-se à análise de indivíduos com SL. Os resultados desta análise revelaram que 5/13 (38%) indivíduos (4, 6, 19, 43 e 45) apresentavam o padrão definido pelo aumento de expressão do transcrito 13a e redução quase total do transcrito *wt*. É de salientar que, independentemente de terem apresentado redução do transcrito *wt*, a maioria dos indivíduos [9/13 (69%)] apresentou sobreexpressão do transcrito 13a, apesar de o indivíduo 44 não apresentar valores significativos devido à variabilidade dos dados obtidos (Figura 23). Desta forma, quer a sobreexpressão do transcrito 13a e redução do *wt*, quer apenas a sobreexpressão do primeiro, foi mais frequente nos indivíduos com SL do que no grupo de indivíduos saudáveis.

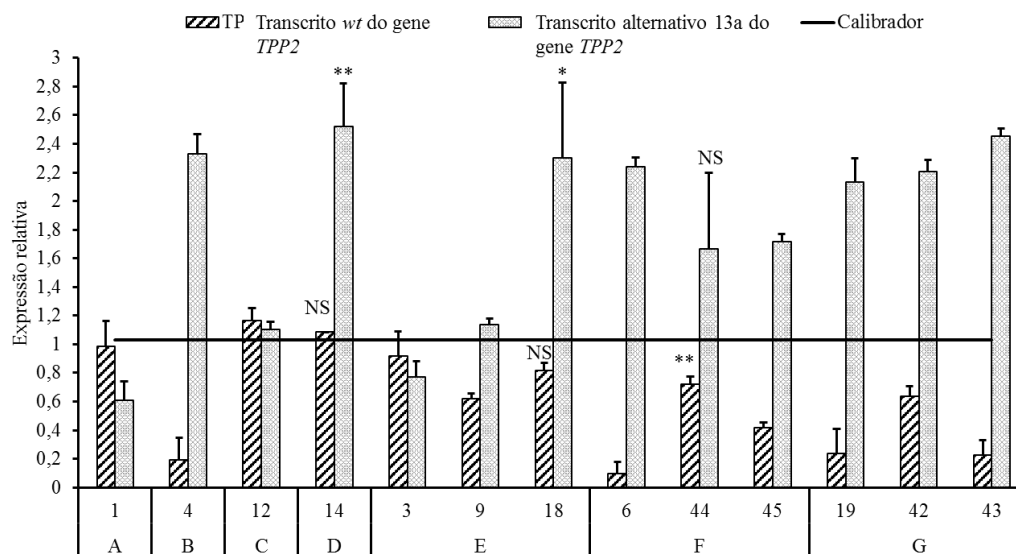


Figura 23. Resultados da expressão relativa do transcrito *wt* e do transcrito 13a do gene *TPP2* em indivíduos com SL. Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). ** $p < 0,01$; * $p < 0,05$; NS – não significativo.

Por fim, para excluir a existência de uma deleção total do gene *TPP2* efectuou-se uma análise de *copy number* utilizando *primers* específicos descritos no anexo I. Estes resultados não revelaram deleção deste gene em nenhum dos indivíduos das famílias analisadas (L55 e L7). Verificou-se apenas que os valores do indivíduo L447 se encontravam ligeiramente aumentados, nomeadamente na região do *TPP2*, no entanto esta amostra parece apresentar também para alguns dos genes anteriores valores de *copy number* ligeiramente mais elevados que as restantes, o que poderá estar associado a uma questão técnica, relacionada com a quantificação de DNA nesta amostra. De acordo com este facto, este indivíduo não é o único indivíduo afectado na família, por isso esta alteração não deverá ter relação com a doença (Figura 24).

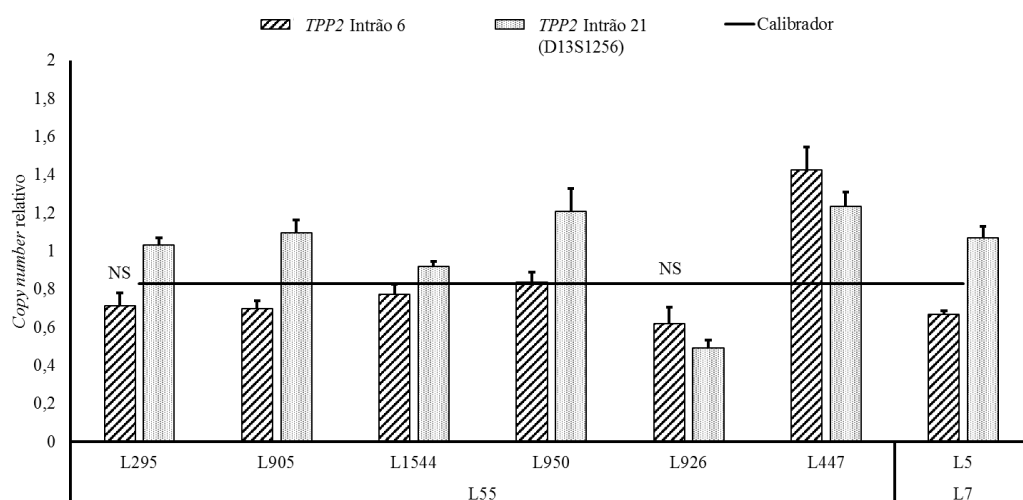


Figura 24. Resultados da análise de *copy number* do gene *TPP2* para a região do intrão 6 e 21 em indivíduos de 2 famílias FCCTX (L55 e L7). Calibrador – Média do *copy number* relativo do grupo controlo. NS – não significativo.

4.2.2. *TEX30*

Num estudo prévio, não foi detectada qualquer mutação pontual no gene *TEX30* (Pereira 2013). No presente estudo foi realizada a análise de grandes deleções/inserções de modo a excluir a existência de outro tipo de alterações neste gene. Esta análise permitiu identificar um transcrito de menor peso molecular com 724pb. Este transcrito foi detectado nos indivíduos L926 da família L55, L5 da família L7 e L452 da família L63 (Figura 25).

A sequenciação deste transcrito ao nível do cDNA identificou uma deleção de 206pb, correspondendo à deleção do exão 5. Através da base de dados *Ensembl* identificou-se esta deleção como característica do transcrito alternativo 4 que compreende 959pb, 4 exões codificantes e uma região 3'UTR maior quando comparado com o transcrito *wt*.

Depois de identificado o transcrito alternativo 4, efectuou-se a pesquisa do mesmo no grupo N e nos indivíduos com SL, tendo sido revelada a presença do transcrito alternativo 4 nos três grupos analisados. Verificou-se assim a presença deste transcrito alternativo com frequências semelhantes nos 3 grupos de indivíduos, em 5/8 (63%) dos indivíduos das famílias FCCTX, em 15/19 (79%) dos indivíduos com SL e em 11/15 (73%) dos indivíduos do grupo N, tendo em consideração a expressão correspondente do transcrito do gene *housekeeping*, *GAPDH* (Figura 25).

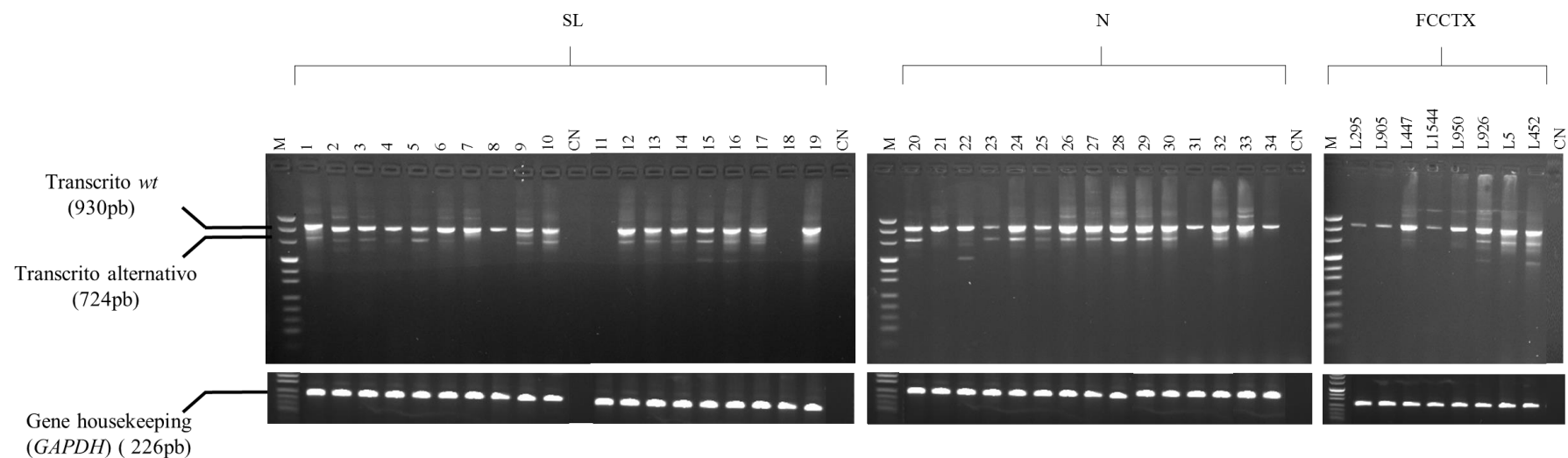


Figura 25. Electroforese em gel de agarose 1,2% (p/v) dos produtos amplificados por PCR para o gene *TEX30* e 2% (p/v) para o gene *housekeeping GAPDH* para os indivíduos de famílias FCCTX, indivíduos com SL e indivíduos saudáveis (N). CN – Controlo negativo; M – Marcador de peso molecular *GeneRuler 50bp DNA Ladder* (Thermo Scientific) (topo) e *GeneRuler 50bp DNA Ladder* (Thermo Scientific) (fundo).

Foi também avaliada a existência de alterações na expressão do gene *TEX30*, estes resultados revelaram que em 29% (2/7) dos indivíduos analisados (L295 e L950) verificava-se uma diminuição de expressão do transcrito *wt* deste gene, no entanto apenas foi possível obter valores significativos para o indivíduo L950, o que, sendo o único membro da família com esta alteração não parece associar-se à doença. Relativamente aos restantes indivíduos, apresentaram valores próximos aos do gene *housekeeping* (Figura 26).

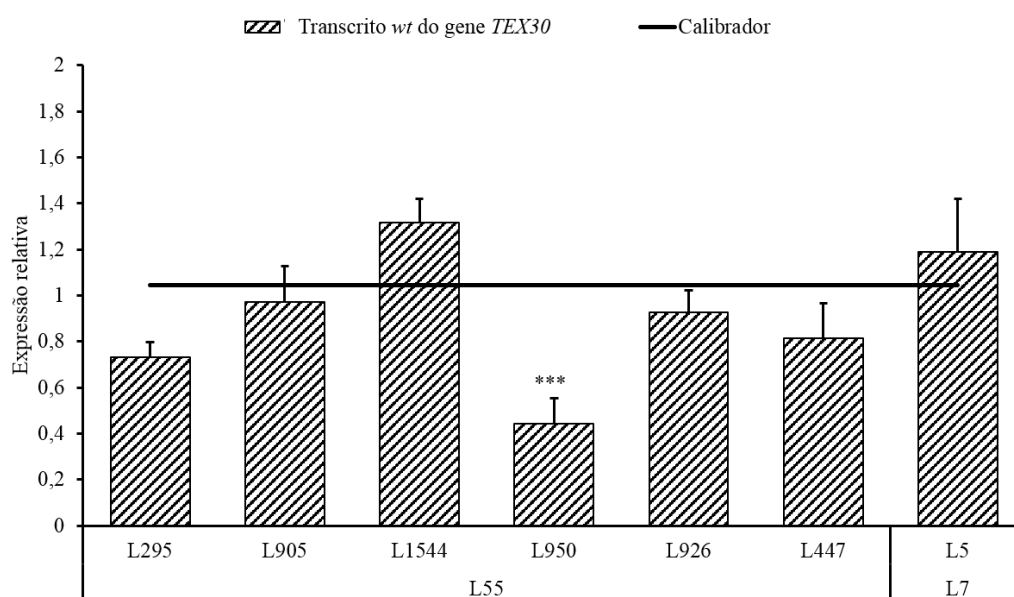


Figura 26. Resultados da expressão relativa do transcrito *wt* do gene *TEX30* para os indivíduos de 2 famílias FCCTX (L55 e L7). Calibrador – Média da expressão relativa dos dois genes *housekeeping* utilizados (*RPS13* e *SDHA*). *** $p < 0,0001$; NS – não significativo.

4.2.3. *SLC10A2*

Num estudo prévio recorrendo a análise *in silico* (softwares *Exonic splicing enhancer finder* (Smith et al. 2006), *Human splicing finder* (Desmet et al. 2009) e *Variant effect predictor* (McLaren et al. 2010)) e análise de segregação com a doença na família (Pereira 2013) foi avaliado o possível carácter patogénico de mutações no gene *SLC10A2*, por uma ou por várias destas análises, as mutações V98I, P290S e A171S revelaram-se possivelmente patogénicas. A mutação V98I revelou segregação com a doença, as mutações P290S e A171S através da análise *in silico* e descrição na literatura (Wong et al. 1995) são sugestivas de possível carácter patogénico.

No presente estudo, foi efectuada uma nova análise *in silico* e análise de segregação para a mutação A171S. No caso das mutações V98I e P290S, tendo em conta a sua baixa frequência na serie FCCTX e na população europeia, foi avaliada a sua frequência na população portuguesa através da análise de indivíduos saudáveis.

Os resultados da análise *in silico* (Tabela 8) permitem constatar que a mutação p.V98I pode ser potencialmente patogénica por criação de novos locais de *splicing*.

Tabela 8. Caracterização das mutações identificadas no gene *SLC10A2*.

Localização	Mutação germinal	Frequência alélica						Variant Effect Predictor		Mutation Taster
		FCCTX		População europeia		Indivíduos saudáveis		SIFT	Polyphen	
Exão 1	c. 292 G>A p.Val98Ile (rs55971546)	A	13%	A	4%	A	5%	Tolerado (0,5)	Benigno (0,317)	Patogénica*
Exão 3	c. 511 G>T p. Ala171Ser (rs188096)	T	13%	T	13%	—	—	Tolerado (1,0)	Benigno (0,001)	Polimorfismo
Exão 5	c. 868 C>T p.Pro290Ser (rs56398830)	T	7%	T	2%	T	3%	Patogénica (0,01)	Patogénica (0,837)	Patogénica* ¹

*Criação de dois novos locais de *splicing*, 2pb a jusante da mutação (*wt*: 0,49 / *mut*: 0,92) e 3pb a montante da mutação (*wt*:0,52 / *mut*: 0,58); *¹ Criação de um novo local de *splicing* 2pb a jusante da mutação (*wt*: 0,53 / *mut*:0,63).

Na série de indivíduos FCCTX a mutação V98I revelou uma frequência alélica de 13% enquanto na população europeia apresentou uma baixa frequência (4%), semelhante à constatada nos indivíduos saudáveis (5%). O *software Mutation taster* revelou a criação de um novo local de *splicing* a 2pb a jusante do local da mutação, com um valor elevado (0,92) em relação ao *wt* (0,49). De forma a esclarecer se a alteração deste valor pode ser relevante, foi analisada a região inicial do intrão 1 (locais críticos de *splicing*) recorrendo ao *software Mutation taster*. Esta análise demonstrou que o local onde é normalmente efectuado o *splicing* do intrão 1 apresenta valores mais baixos do que os valores obtidos pela mutação V98I (0,92), o que pode implicar alguma alteração do *splicing* por parte desta mutação. Os locais analisados foram a última base do exão 1 (377pb – *wt*: 0,72), a primeira base do

intrão 1 (377pb+1 – wt: 0,49) e a quinta base do intrão 1 (377pb+5 – wt: 0,78), estes valores sugerem a eventual patogenicidade desta mutação.

A mutação P290S tinha sido detectada com uma frequência de 7% nas famílias de FCCTX. No presente estudo observou-se uma frequência de 3% num grupo de 47 indivíduos saudáveis, a qual é semelhante à descrita na população europeia (2%). Para além da baixa frequência constatada, a análise *in silico* revelou esta mutação como possuindo possivelmente carácter patogénico.

Em relação à mutação A171S, apesar de se encontrar descrita como um polimorfismo (rs188096), com uma frequência na população europeia de 13%, existia a hipótese desta mutação poder contribuir para o fenótipo desta família (L84) em associação com a variante de susceptibilidade E1317Q previamente identificada na referida família. No entanto, a mutação foi identificada em heterozigotia em 2 indivíduos afectados e 1 indivíduo não afectado, enquanto noutro indivíduo não afectado a mutação foi identificada em homozigotia.

4.3. Análise de *copy number* na região 13q32-33 em tumores de indivíduos FCCTX

De modo a esclarecer os resultados de LOH obtidos por Pereira em 2013, no âmbito da análise por *genescan* da região 13q32-33 (Figura 27), efectuou-se uma análise de *copy number* em três regiões que compreendem os genes *DOCK9* (próximo do marcador D13S1298) e *TPP2* (D13S1256 – intragénico) (Figura 28), utilizando os *primers* específicos descritos no anexo I. Na figura 29 encontram-se representados os resultados da análise de *copy number* de uma forma esquemática.

Os resultados desta análise revelaram que o caso 437/438 apresenta uma amplificação de ambos os genes, tanto no carcinoma como na mucosa normal, mas sobretudo no primeiro, no entanto a análise de LOH revelou-se como normal para os marcadores D13S1298 e D13S1266 e não informativo para o marcador D13S1256.

O caso 1541/1542 parece apresentar um ganho em ambas as regiões do gene *TPP2*, no caso da mucosa normal, e revelou uma análise de LOH normal para o marcador D13S1256. Estes casos que apresentam ganho/amplificação em apenas um ou em ambos os genes analisados, com análise de LOH normal, podem tratar-se de casos com ganho/amplificação dos 2 alelos, alteração que não é detectada por análise de LOH, uma vez que esta detecta uma perda de heterozigotia, ou seja, uma deleção ou um ganho de um alelo em relação ao outro.

No caso 595/596 verificou-se que existe uma amplificação no carcinoma, quer na região do *DOCK9* quer na região do *TPP2*, a análise de LOH revelou-se como não informativa no marcador D13S1266 e inconclusiva para o D13S1256, podendo tratar-se de LOH ou de MSI. No marcador D13S1298 não foi possível apresentar resultado mas no marcador anterior a este apresenta LOH. Desta forma, parece confirmar-se a amplificação no carcinoma em relação ao normal. Relativamente ao caso 1773/1774 verificou-se que parece existir uma amplificação na região do intrão 21 do gene *TPP2*, no adenoma. A análise de LOH revelou MSI no adenoma e perda de um alelo na mucosa normal. Não revelou LOH para os restantes marcadores mas apresenta ligeira MSI em alguns marcadores em 13q33. Uma vez que a MSI impossibilita o cálculo de LOH, o adenoma poderá apresentar ganho em relação à mucosa normal.

O caso 1777 apresentou ganho/amplificação na região do intrão 21 do gene *TPP2*, no adenoma, no entanto devido à discrepância de valores não foi significativo. A análise de LOH desta amostra revelou-se com perda de heterozigotia nesta região que corresponde a este ganho detectado. É de notar que estas amostras (595/596, 1773/1774 e 1777) apresentam ainda MSI em 1 ou mais marcadores da região analisada do cromossoma 13, apesar destes adenomas serem MSS.

Na análise realizada, os casos 991 e 1791/992 e 993 apresentaram valores normais no adenoma e na mucosa normal, mas a análise de LOH revelou perda de heterozigotia na região do marcador D13S1256, prolongando-se até ao marcador D13S1298 no caso da amostra 991. A LOH na ausência de variação do número de cópias poderá implicar LOH com manutenção do número de cópias (*copy-neutral*) que resulta de mecanismos de recombinação com duplicação de um dos alelos e perda do outro. Os tumores, neste caso adenomas, poderão apresentar um único alelo duplicado sendo o número de cópias mantido. É de notar que também os casos 1791 e 993 apresentam MSI em 1 ou mais marcadores, apesar de serem MSS, de acordo com a análise dos marcadores de Bethesda, os quais são altamente sensíveis para a instabilidade de alto-grau.

No caso dos dois adenomas 2861A1 e 2862A2 e respectiva mucosa normal, verificou-se amplificação nas duas lesões e também na mucosa normal. A análise de LOH tinha revelado, para o marcador D13S1256, perda total de um alelo para a mucosa normal e perda total do outro alelo para os adenomas. Para o marcador D13S1266 verificava-se MSI. A amplificação observada na análise de *copy-number*

pode estar associada a recombinação com duplicação de um dos alelos e eliminação do outro alelo, no entanto, quer os resultados de LOH quer de *copy-number* necessitam de confirmação.

No caso do 1871/1872 parece verificar-se um ligeiro ganho no carcinoma em relação à mucosa normal, o que terá de ser confirmado devido à ausência de LOH nesta região.

O caso 569/570 parece apresentar uma deleção na região do intrão 21 do gene *TPP2*, no carcinoma, e em ambas as regiões do mesmo gene, na mucosa normal. No entanto, como a análise de LOH era normal será importante repetir este caso para confirmar este resultado.

No caso do 947/948 parece verificar-se um ganho apenas da região do gene *DOCK9* no adenoma, no entanto a mucosa normal apresentou amplificação para os dois genes. No caso deste adenoma, apesar de se verificar LOH na região do *DOCK9* e *TPP2*, existe dúvida quanto ao resultado do D13S1256, em particular, no que diz respeito a uma ligeira MSI. A LOH observada poderá resultar de deleção no adenoma 947 em relação à mucosa normal.

O caso 459/460 apresentou valores normais no tumor e na mucosa normal. Sendo que não revelou LOH, pode-se constatar que este tumor não tem alteração no número de cópias destes dois genes.

L	DNA		D13S795	D13S167	D13S1278	D13S1280	D13S132	D13S1298	D13S225	D13S174	D13S1266	D13S1256	D13S158
7	437	438		NI		NI						NI	
24	947	948											
	459	460							NI				
55	991	992									NI		
	1791			NI							NI		
	993	L905		NI							NI		
	1773	1774	NI	NI	NI	NI				NI	NI		
	1777	L950											NI
	1871	1872									NI		
	2861-A1	2862											
	2861-A2												
63	569	570			NI				NI				NI
	595	596		NI	NI				NI		NI		
P2209	1541	1542								NI			

Figura 27. Resultados da análise de LOH para os marcadores de microssatélites D13S795 a D13S158. Verde – Normal; Vermelho – Perda de heterozigotia; Azul – Instabilidade de microssatélite; Amarelo – Inconclusivo; NI – Não informativo; Sem preenchimento – Sem resultado disponível.

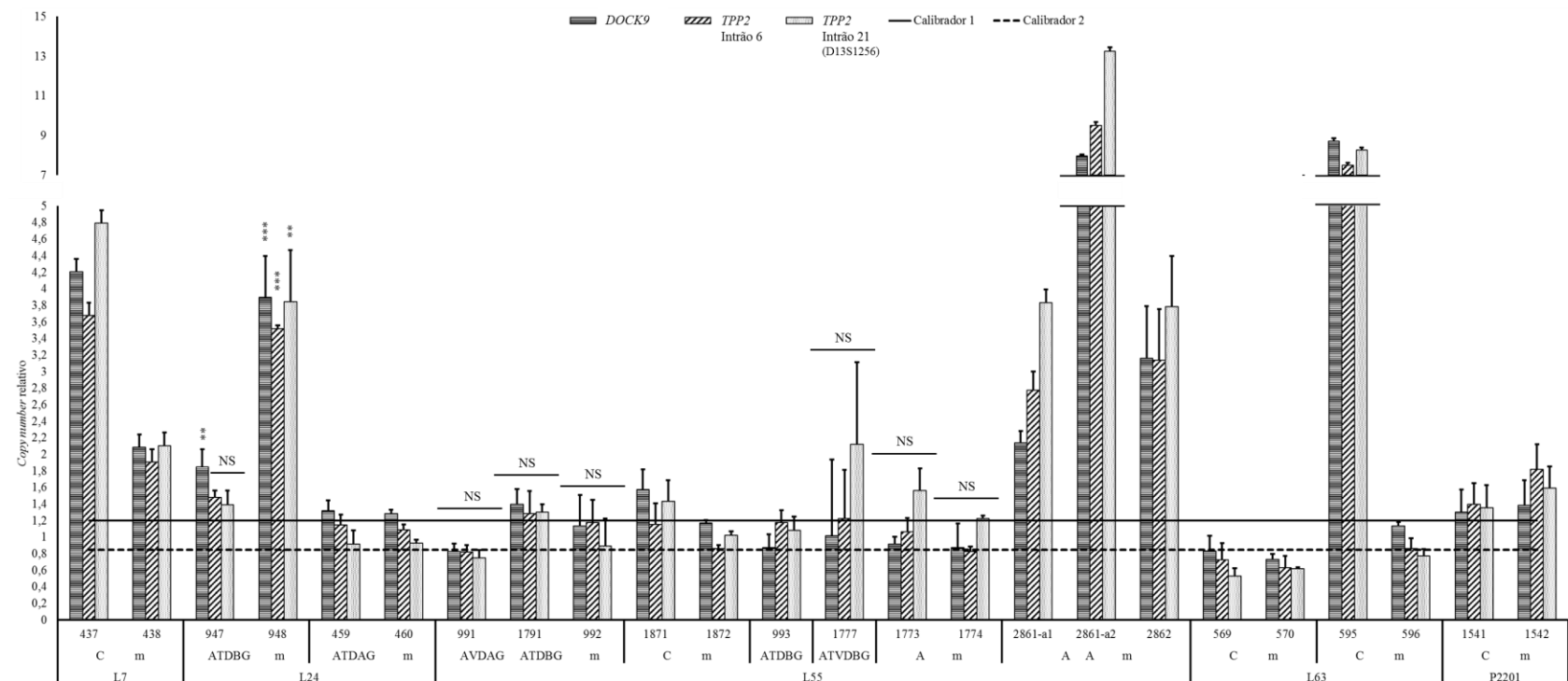


Figura 28. Resultados de *copy number* dos genes *DOCK9* e *TPP2* regiões do intrão 6 e 21 para os tumores e mucosas normais dos indivíduos das famílias FCCTX. C – carcinoma; A – adenoma; m – mucosa normal; ATDAG – adenoma tubular com displasia de alto grau; ATDBG - adenoma tubular displasia baixo grau; AVDAG – adenoma viloso com displasia alto grau; ATVDBG – adenoma túbulo - viloso com displasia baixo grau; Calibrador 1 – Valor máximo da média obtida no grupo controle; Calibrador 2 – Valor mínimo da média obtida no grupo controle. *** $p < 0,0001$; ** $p < 0,001$; NS – Não significativo.

Família	DNA	Histologia	CIN	D13S1298	—	D13S1256
				<i>DOCK9</i>	<i>TPP2</i> intrão 6	<i>TPP2</i> intrão 21
L7	437	C	+	4,21	3,68	4,79
	438	m		2,09	1,91	2,10
L24	947	ATDBG	+	1,85	1,48	1,39
	948	m		3,89	3,52	3,85
	459	ATDAG	+	1,32	1,15	0,92
	460	m		1,28	1,09	0,93
L55	991	ATDAG	+	0,82	0,82	0,75
	1791	ATDBG		1,40	1,29	1,30
	992	m		1,14	1,18	0,89
	1871	C	+	1,58	1,15	1,43
	1872	m		1,17	0,86	1,02
	993	ATDBG	+	0,87	1,18	1,08
	1777	ATVDBG	+	1,02	1,23	2,12
	1773	A	+	0,91	1,06	1,56
	1774	m		0,88	0,83	1,23
	2861-a1	ATVDBG	NE	2,14	2,77	3,83
	2861-a2	ATVDBG		7,97	9,50	13,25
	2862	m		3,16	3,14	3,79
L63	569	C	+	0,83	0,73	0,53
	570	m		0,73	0,63	0,62
	595	C	+	8,71	7,50	8,26
	596	m		1,13	0,86	0,77
P2201	1541	C	-	1,30	1,40	1,36
	1542	m		1,38	1,82	1,59
-	2616	m	NE	1,10	1,31	1,20
	2594	m		0,91	0,78	0,85

Figura 29. Resultados da análise de *copy number* da região 13q32-33 através da análise dos genes *DOCK9* e *TPP2*. C – carcinoma; A – adenoma; m – mucosa normal; ATDAG – adenoma tubular com displasia de alto grau; ATDBG - adenoma tubular displasia baixo grau; AVDAG – adenoma viloso com displasia alto grau; ATVDBG – adenoma túbulo - viloso com displasia baixo grau; Verde – Amplificação do número de cópias de DNA; Vermelho – Diminuição do número de cópias de DNA; NE – Não efectuado.

5. Discussão

As famílias FCCTX caracterizam-se pelo preenchimento dos CA, ausência de mutações germinais nos genes de reparação de erros de DNA do tipo *mismatch* e tumores com ausência de instabilidade de microssatélite. Contudo a causa molecular deste síndrome ainda não é conhecida, incentivando deste modo o estudo de genes que possam estar associados à susceptibilidade para o desenvolvimento de CCR nas famílias FCCTX.

Em estudos prévios, o grupo de Gastrenterologia da Unidade de Investigação de Patobiologia Molecular do IPOFLG, EPE de Lisboa, demonstrou a presença de *linkage* com a doença numa família FCCTX, nas regiões 13q e 21q, que posteriormente foram delimitadas recorrendo à análise de perdas de heterozigotia utilizando marcadores de microssatélite. Estas regiões foram delimitadas a uma região mínima de LOH de 0,87Mb em 13q33 e 1,3Mb em 21q11, sendo constatado ainda que LOH em 13q ocorre maioritariamente em adenomas e LOH em 21q em carcinomas, sugerindo uma associação entre: iniciação tumoral e alterações na região 13q, e progressão tumoral e alterações na região 21q.

Envolvimento do gene candidato *BMPRIA* como gene de susceptibilidade para o FCCTX

A análise de mutações pontuais do gene *BMPRIA* não revelou nenhuma das mutações encontradas no estudo de Nieminen realizado em 2011 (Nieminen et al. 2011), no entanto identificaram-se as mutações p.Asp380Tyr ec.1599+847 G>A com um possível carácter patogénico. As restantes mutações, devido à análise *in silico* não prever uma eventual patogenicidade e/ou devido a uma elevada frequência na série estudada e também na população europeia, foram excluídas como possíveis mutações patogénicas como o descrito anteriormente (4.1.1), o que sugere tratarem-se de variantes polimórficas sem carácter patogénico. A mutação *missense* A380T foi revelada como patogénica no *software Mutation taster*, devido à criação de novo local de *splicing*, no entanto a frequência alélica relativamente elevada na população europeia (7%), sugere tratar-se de uma variante polimórfica, não sendo provável a sua associação com o FCCTX. A mutação c.1599+847 G>A pode não estar documentada, nas bases de dados analisadas, devido à sua localização, sendo que a

região 3'UTR não é analisada primordialmente em estudos de análise mutacional. A sua localização diminui também a probabilidade de se tratar de uma mutação com carácter patogénico pois a criação de um local de *splicing* nesta região não deverá, à partida, modificar o *splicing* do gene, não afectando a síntese da proteína. No entanto, para ambas as mutações, os resultados da análise *in silico* em associação à sua baixa frequência alélica nas famílias FCCTX (3%) não permite excluir completamente a possibilidade de contribuir para a susceptibilidade deste síndrome, por exemplo em associação com outra mutação, sendo necessário efectuar-se estudos de segregação com a doença na família onde foi detectada.

É de salientar ainda a mutação p.Pro2Thr, a qual segundo Kaczmarczyk e Böttcher, encontra-se descrita com carácter patogénico devido a interferir com o transporte pós-traducional da proteína recém sintetizada (Kaczmarczyk et al. 2013; Böttcher et al. 2009), no entanto a análise *in silico* realizada e a sua elevada frequência na população europeia (25%) e nas famílias FCCTX (24%) não sugere contribuir para o carácter patogénico deste síndrome. Relativamente à mutação c.1599+1497 T>C, que revelou uma frequência alélica baixa nas famílias FCCTX (12%) em relação à frequência na população europeia (27%), esta variação parece ser populacional tendo em conta as frequências das populações asiática (9%) e africana (14%). Relativamente às restantes mutações identificadas, estas apresentam frequências alélicas elevadas na população europeia (11%-28%) o que sugere tratar-se de variantes polimórficas sem carácter patogénico.

Envolvimento de genes candidatos na região 21q11 na susceptibilidade para o FCCTX

A análise de mutações pontuais nos genes *HSPA13* e *SAMSNI* não revelou mutações com carácter patogénico quer pela elevada frequência na população estudada quer pela análise *in silico* efectuada. Relativamente ao gene *HSPA13*, não foi identificado nenhum dos polimorfismos presentes no estudo em carcinoma gástrico realizado por Aoki em 2005 (Aoki et al. 2005) e não se encontram descritas mutações germinais patogénicas neste gene. A mutação c.749-56 G>A, no gene *HSPA13*, não se encontra documentada nas bases de dados analisadas, no entanto ao tratar-se de uma mutação intrónica, a 56pb do exão mais próximo, pode não ter sido

detectada em estudos de análise mutacional. Apesar desta mutação apresentar uma frequência alélica baixa nas famílias FCCTX (3%), a análise *in silico* exclui o seu possível carácter patogénico. Relativamente ao gene *SAMSN1*, a elevada frequência das mutações detectadas e o facto da análise *in silico* não prever qualquer carácter patogénico, sugere tratarem-se de variantes polimórficas sem carácter patogénico. Segundo a base de dados COSMIC (Forbes et al. 2011), a percentagem de mutações somáticas destes dois genes é muito baixa, sendo predominante na região do cego, no caso *HSPA13* (2,97%), e na região do recto, no caso do *SAMSN1* (3,82%), no entanto esta informação por si só não excluía o seu envolvimento a nível germinal sabendo-se que o gene *SMAD4* apresenta uma baixa frequência de mutações somáticas em cólon (13,58%) mas no entanto está associado ao síndrome de polipose juvenil (Howe 2004).

Assim, o presente estudo sugere que os genes *HSPA13* e *SAMSN1* não parecem estar envolvidos na susceptibilidade para o FCCTX a nível de mutações pontuais germinais, sendo que apesar de estarem documentadas mutações somáticas (*missense*) com carácter patogénico estas não implicam o envolvimento germinal destes genes.

Adicionalmente, foi avaliado o envolvimento de possíveis alterações de expressão nestes genes, ao nível germinal, na susceptibilidade para o FCCTX. Verificaram-se algumas alterações de expressão nestes dois genes, nomeadamente o aumento de expressão do transcrito *wt* nos indivíduos da família L55 (L1544 e L447). No entanto, como este se verificava para ambos os genes, *HSPA13* e *SAMSN1*, teria que resultar de alguma alteração que envolvesse os dois genes em simultâneo como uma duplicação dessa região cromossómica. No entanto a análise de *copy number* efectuada para ambos os genes revelou apenas para o indivíduo L447 um ligeiro ganho do gene *SAMSN1*, sendo necessário repetir esta análise para confirmação. Verificando-se que não existe ganho/amplificação que justifique o aumento de expressão destes dois genes independentes e sabendo-se que estes dois indivíduos não são os únicos afectados na família L55, esta alteração de expressão não deverá contribuir para a susceptibilidade para este síndrome, nestes casos, podendo tratar-se de alguma alteração a nível fisiológico. Como exemplo desta última, pode ser o envolvimento numa resposta inflamatória, que pode ou não ser desencadeada por um processo de tumorigénese. Ambos os genes podem estar envolvidos em processos inflamatórios, a *HSPA13* relacionada com a apoptose

celular mediada pela via TRAIL, como sugerido por Grizenkova em 2012 (Grizenkova et al. 2012). A sobreexpressão do gene *SAMSN1* está relacionada com a activação dos linfócitos B (Zhu et al. 2004) e segundo o estudo de Wang em 2010, este gene intervém na via de inibição PIR-B inibindo a proteína Lyn e regulando negativamente os linfócitos B (Wang et al. 2010).

No estudo de Yamada em 2008, realizado em linhas celulares de carcinoma do pulmão verificou-se uma diminuição de expressão do gene *SAMSN1*, sendo posteriormente proposto como gene supressor de tumor em cancro do pulmão (Yamada et al. 2008). Nas famílias de FCCTX parece existir um indivíduo (L295) da família L55 com uma diminuição de expressão do transcrito deste gene, sendo necessário repetir esta análise para confirmar esta alteração. A análise de *copy number* deste gene revelou uma possível deleção em dois indivíduos (L926 e L5) das famílias L55 e L5, sendo necessário também repetir a análise para confirmação.

Deste modo, os genes *HSPA13* e *SAMSN1* não parecem estar envolvidos na susceptibilidade para o FCCTX, apesar de ser necessário repetir as análises de expressão e de *copy number* de forma a garantir que as alterações não se tratam de erros associados à técnica. Na região do cromossoma 21 delimitada pelos estudos de *linkage* e de LOH, encontram-se vários pseudogenes que limitam a possibilidade de genes candidatos para estudo, sugerindo-se assim concluir os estudo do gene *NRIP1*, previamente estudado a nível de mutações pontuais, nomeadamente efectuando uma análise de expressão e de *copy number*. Recentemente, o estudo de expressão no gene *NRIP1* de Lapierre em 2014, revelou que este gene inibe a proliferação de células cancerígenas de colon *in vitro* e parece estimular a transcrição do gene *APC*, interveniente na via Wnt/ β -catenina, que mantém a homeostase do epitélio cólico. Este estudo sugere ainda um papel de gene supressor em cancro do colon (Lapierre et al. 2014). Recorrendo à base de dados COSMIC verifica-se ainda que este gene apresenta uma frequência de mutações somáticas (2,93%) e alterações de expressão (7,55%) superiores aos genes *HSPA13* e *SAMSN1*.

Envolvimento de genes candidatos na região 13q32-33 na susceptibilidade para o FCCTX

Num estudo prévio foi identificado um transcrito alternativo no gene *TPP2* com uma inserção de 72pb que à data do presente estudo se verificou corresponder à inserção de um possível exão entre os exões 13 e 14 (exão 13a). No presente estudo foi detectado, por PCR quantitativo, um aumento de expressão do transcrito 13a nas famílias de FCCTX (71%), em indivíduos com SL (69%) e num grupo de indivíduos saudáveis (35%). Por outro lado, juntamente com o aumento de expressão do transcrito 13a, foi detectada também em alguns casos redução de expressão do transcrito *wt*, mais frequente também nos indivíduos FCCTX e SL (43% e 38%, respectivamente) do que nos indivíduos saudáveis (18%).

A função deste gene não está esclarecida e por isso não se sabe com certeza qual o papel deste transcrito alternativo, não documentado, nem do próprio transcrito *wt*. No entanto, sendo o aumento de expressão do transcrito 13a frequente nos três grupos analisados, e estando a proteína TPP2 envolvida na resposta imune, poderá tratar-se de uma alteração associada a algum processo fisiológico comum, não sendo, neste momento possível discriminar qual o papel do transcrito alternativo em relação ao transcrito *wt*. De facto, encontra-se descrita a existência de uma relação entre a expressão do *TPP2* e a apresentação antigénica envolvida na resposta imunitária (Schnurr et al. 2009).

Apesar da elevada frequência do aumento de expressão do transcrito 13a nos indivíduos FCCTX, esta poderá ser influenciada pelo facto de terem sido analisados 6 indivíduos da mesma família. No entanto, nos casos de SL também se verificou um aumento da expressão do transcrito 13a e redução do *wt*, com uma frequência superior à observada nos indivíduos saudáveis. Nos casos de SL encontra-se descrito que devido à deficiência do sistema MMR que origina mutações *frameshift* nas regiões de microsatélite, pode verificar-se a indução da tradução de *immunogenic frameshift peptides* (FSP). Estes são apresentados como antígenos aos linfócitos T citotóxicos, estimulando uma resposta imunitária contra o tumor, o que poderá estar associado aos infiltrados inflamatórios característicos de tumores de SL (Doeberitz & Kloor 2013). Estes são apresentados aos linfócitos T citotóxicos através complexo principal de histocompatibilidade I (MHC-I). A protease TPP2 tem sido descrita como estando envolvida no processo mediado pelo MHC-1, sugerindo que as

alterações de expressão constatadas com maior frequência nos indivíduos com SL pode relacionar-se com este processo inflamatório (Schnurr et al. 2009).

Relativamente ao gene *TEX30* que não revelou qualquer mutação germinal com carácter patogénico em estudos anteriores, observou-se na análise de grandes deleções um transcrito alternativo com expressão semelhante nos três grupos analisados, FCCTX, SL e indivíduos saudáveis. A análise de expressão revelou apenas um indivíduo com redução de expressão deste gene (L950), mas não sendo o único indivíduo afectado da família L55, esta alteração não deverá estar envolvida na susceptibilidade para o FCCTX. Os estudos realizados com este gene sugerem assim que o mesmo não está envolvido na susceptibilidade para o FCCTX.

No que diz respeito ao gene *SLC10A2*, a análise de segregação da mutação A171S revelou que a presença desta mutação em homozigotia num indivíduo não afectado parece, de facto, excluir a possibilidade de contribuir para o fenótipo desta família, em conjunto com a mutação E1317Q previamente identificada no gene *APC*. Em relação à mutação P290S, a frequência relativamente baixa na população europeia (2%) e no grupo de indivíduos saudáveis analisados no presente estudo (3%), aliada à análise *in silico* não são suficientes para excluir o seu envolvimento, pelo que seria importante realizar uma análise de segregação da mutação com a doença na família.

A mutação V98I, de acordo com a análise *in silico*, conduz à criação de um local de *splicing* 2pb a jusante do local da mutação e depois de ser identificado o possível local de *splicing wt* e constatando-se que este local tem valores de probabilidade de *splicing* mais baixos do que o do local criado pela mutação, pode-se sugerir um eventual efeito patogénico para esta mutação, podendo a mesma originar uma proteína que perde parte do exão 1. Será necessário estudar esta possível alteração ao nível do RNA, de modo a avaliar se existe uma deleção parcial do referido exão. Este facto é reforçado, tendo em conta que esta mutação parece segregar com a doença na família, apesar desta avaliação ainda não ser conclusiva.

Envolvimento de LOH e alterações de *copy number* em 13q32-33 na tumorigénese associada ao FCCTX

A análise de *copy number* dos tumores FCCTX, na região 13q32-33, permitiu esclarecer quanto à natureza da LOH na maioria dos tumores que apresentavam esta alteração. De facto, na maioria dos casos, a LOH correspondeu a ganhos/amplificações ou então eventualmente a LOH com manutenção do número de cópias (*copy-neutral*),

No caso das amostras de adenoma e carcinoma da família L55, alguns adenomas que apresentavam LOH revelaram manutenção do número de cópias (2 ATDBG e 1 ATDAG), o que poderá ser explicado por *copy-neutral* LOH devido a recombinação ou não disjunção. No caso do ATVDBG (1777) a LOH foi confirmada como ganho. Os outros dois ATVDBG (2861A1 e 2861A2) também apresentaram ganho/amplificação, com a particularidade de a mucosa normal respectiva apresentar já ganho/amplificação, o que revela ser um evento bastante precoce. É de notar que esta mucosa normal apresentava perda total de um dos alelos para o marcador D13S1256, em relação aos alelos observados na linha germinal (sangue) (resultados não apresentados neste estudo), o que poderá sugerir que possa ter ocorrido duplicação/amplificação de um dos alelos, enquanto o segundo alelo era eliminado, num mecanismo com uma génese semelhante ao *copy-neutral*. De acordo, também foram observados ganhos nesta região noutras mucosas normais de indivíduos de outras famílias (438, 948 e 1542), o que sugere que este poderá ser um mecanismo frequente na tumorigénese associada ao FCCTX. Também frequente nestes tumores em particular é a existência de uma ligeira MSI em um ou mais marcadores de microsatélites, apesar de classificados como MSS para os marcadores de Bethesda, o que evidência também uma assinatura molecular específica. Estes ganhos/amplificações e ligeira MSI muito precoces (mucosa normal e adenomas) sugere um elevado grau de instabilidade genómica, o que pode ser importante em termos de selecção de genes candidatos para a susceptibilidade para este síndrome.

O único carcinoma da família L55 (1871), não apresentou LOH mas, pelo facto de apresentar um ligeiro aumento de *copy-number*, deverá ser repetido. Desta forma parecem observar-se diferentes resultados para a mesma família, os quais parecem estar associados a diferentes histologias.

Também foram detectados ganhos/amplificações em casos que não apresentavam LOH, no caso da mesma família, como no caso do carcinoma 595, que deverá por isso corresponder a amplificação desta região nos dois alelos, alteração esta que poderá ser compatível com a amplificação frequente em 13q observada em carcinomas colorectais (Angelis et al. 1999). Aparentemente, os ganhos/amplificações em mucosa normal e adenoma não deverão corresponder a esta amplificação em 13q, a qual se encontra associada à progressão tumoral. Adicionalmente, neste grupo de tumores FCCTX, estes ganhos/amplificações foram mais frequentes em adenoma do que em carcinoma, o que pode sugerir que podem não conferir vantagem selectiva para a progressão tumoral.

Apenas um único caso (569/570) apresentou deleção no carcinoma e respectiva mucosa normal nas regiões dos genes *DOCK9* e *TPP2*. No entanto, devido à redução do valor de *copy-number* não ser significativa, este caso terá de ser repetido, principalmente por não apresentar LOH. A confirmar-se uma deleção poderá ser explicada por deleção em ambos os alelos.

6. Conclusão

Aproximadamente metade das famílias de cancro do cólon e recto hereditário não associado a polipose (HNPCC) (que preenchem os CA) não apresentam mutação germinal nos genes MMR e os tumores são MSS. Estes casos são por isso designados por cancro do cólon e recto do tipo X (FCCTX), cuja causa molecular é ainda desconhecida.

No presente estudo foi avaliado o envolvimento de genes específicos na susceptibilidade para o FCCTX. Com base nos resultados de Nieminen e colaboradores foi realizada uma análise de mutações pontuais germinais no gene *BMPRIA*. Na região 21q11, estudou-se o envolvimento dos genes *HSPA13* e *SAMSNI* na susceptibilidade para esta patologia, através da pesquisa de mutações pontuais, grandes deleções/inserções e alterações de expressão, ao nível germinal. No caso da região 13q32-33, o envolvimento de alguns genes candidatos, já estudados previamente, não estava excluído por completo, nomeadamente o gene *TEX30* ao nível da alteração de expressão ou de grandes deleções/inserções e levantou-se a hipótese de algumas alterações nos genes *TPP2* (expressão de um transcrito alternativo) e *SLC10A2* (mutações p. V98I, p. P290S e p.A171S) poderem estar associadas a este síndrome. Ainda na região 13q32-33 foi realizada uma análise de *copy number* em tumores FCCTX de forma a esclarecer a natureza da LOH, detectada num estudo prévio.

Quanto ao envolvimento do gene *BMPRIA*, a pesquisa de mutações germinais neste gene, não revelou quaisquer mutações com potencial evidentemente patogénico nem foram identificadas as mutações do estudo de Nieminen em nenhum dos indivíduos analisados. No entanto, foram identificadas 2 mutações (A380T e c.1599+847 G>A), para as quais ainda não é possível excluir por completo a sua possível patogenicidade, sendo por isso recomendada uma análise de segregação com a doença na família respectiva. O estudo dos genes *HSPA13* e *SAMSNI* localizados na região 21q11 não revelou qualquer alteração com carácter patogénico que possa ser associada com a susceptibilidade para o FCCTX. Foram apenas detectadas alterações de expressão nos genes *HSPA13* e *SAMSNI* que sugerem estar relacionadas com uma eventual resposta inflamatória.

No que diz respeito aos genes localizados em 13q32-33, a conclusão da análise do gene *TEX30* revelou um transcrito alternativo presente com um padrão e

frequência semelhante em famílias FCCTX e em indivíduos saudáveis e por isso não está associado a este síndrome. Relativamente ao gene *SLC10A2* conclui-se que a mutação A171S não contribui para o fenótipo da família e a mutação P290S, apesar de não ser potencialmente patogénica, deverá ser efectuada análise de segregação da mutação com a doença na família, de modo a poder excluí-la. A mutação V98I revelou na análise *in silico* e de segregação indícios de carácter patogénico sendo necessário avaliar a sua relevância, ao nível da análise de expressão de eventuais transcritos aberrantes. A análise de expressão do gene *TPP2* revelou um aumento de expressão de um transcrito específico (13a) deste gene maioritariamente em indivíduos FCCTX e indivíduos com SL, o que poderá sugerir uma eventual relação com a tumorigénese nestas duas entidades moleculares ou com a resposta imune. Esta última hipótese assume maior relevância uma vez que este transcrito também é sobreexpresso em alguns indivíduos saudáveis e tendo em conta a função da proteína TPP2 na apresentação de antígenos mediada pelo MHC-1. Assim, fica por esclarecer o papel do transcrito 13a. A análise de *copy number* na região 13q32-q33 nos tumores FCCTX permitiu esclarecer que parece existir mais do que um fenómeno molecular na base das perdas de heterozigotia detectadas na análise de LOH, sendo necessário efectuar mais estudos de forma a perceber se algum dos fenómenos se encontra directamente relacionado com a tumorigénese do FCCTX. No que diz respeito à LOH frequente em 13q32-q33, a identificação de ganhos/amplificações frequentes em lesões precoces sugere uma assinatura molecular específica nestes tumores, aliada a um elevado grau de instabilidade genómica.

Em resumo, este trabalho exclui os genes localizados em 21q11, *HSPA13* e *SAMSNI*, e em 13q32-33, *TEX30*, como possíveis genes candidatos para o FCCTX, no entanto ainda não é possível excluir o gene *SLC10A2* nem o gene *TPP2*. No que diz respeito ao gene *BMPRI1A*, mutações neste gene não parecem contribuir com frequência para a susceptibilidade para o FCCTX, podendo mesmo ser raras.

Foi ainda identificada uma assinatura molecular específica em lesões precoces FCCTX, com ganhos/amplificações frequentes, aliada a um elevado grau de instabilidade genómica.

7. Referências bibliográficas

- Angelis, P. et al., (1999) Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumour DNA ploidy, genotypes and phenotypes. *British journal of cancer*, 80(3), pp.526–35.
- Aoki, M. et al., (2005) A genetic variant in the gene encoding the stress70 protein chaperone family member STCH is associated with gastric cancer in the Japanese population. *Biochemical and biophysical research communications*, 335(2), pp.566–74.
- Applied Biosystems, (2009) *DNA Sequencing by Capillary Electrophoresis DNA Sequencing by Capillary Electrophoresis, Applied Biosystems chemistry guide* Second (Ed.), Applied Biosystems.
- Attisano, L. & Wrana, J.L., (2002) Signal transduction by the TGF-beta superfamily. *Science*, 296(5573), pp.1646–7.
- Beck, S.E. et al., (2006) Bone morphogenetic protein signaling and growth suppression in colon cancer. *American journal of physiology. Gastrointestinal and liver physiology*, 291(1), pp.135–145.
- Belo, H., (2010) “*Papel do gene APC e identificação de regiões cromossômicas envolvidas no síndrome familiar de cancro do cólon e recto do tipo X.*”, Tese de mestrado em Biologia molecular em saúde, Escola Superior de Saúde Egas Moniz, Portugal.
- Boland, (2005) Evolution of the nomenclature for the hereditary colorectal cancer syndromes. *Fam Cancer*, 4(3), pp.211–218.
- Boland & Goel, (2011) Microsatellite Instability in Colorectal Cancer. *National Institute of Health*, 138(6), pp.2073–2087.
- Böttcher, Y. et al., (2009) Adipose Tissue Expression and Genetic Variants of the. *DIABETES*, 58, pp.2119–2128.
- Cao, X. & Chen, D., (2009) The BMP signaling and in vivo bone formation. , 357(1), pp.1–8.
- Claudio, J.O. et al., (2001) HAC1 encodes a novel SH3-SAM adaptor protein differentially expressed in normal and malignant hematopoietic cells. *Oncogene*, 20, pp.5373–5377.
- Desmet, F.-O. et al., (2009) Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic acids research*, 37(9), pp.1–14.
- Diep, C.B. et al., (2006) The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. *Genes, chromosomes & cancer*, 45(1), pp.31–41.

- Doeberitz, M. & Kloor, M., (2013) Towards a vaccine to prevent cancer in Lynch syndrome patients. *Familial cancer*, 12(2), pp.307–12.
- Duan, S. et al., (2012) FBXO11 targets BCL6 for degradation and is inactivated in diffuse large B-cell lymphomas. *Nature*, 481(7379), pp.90–3.
- Ensembl, (2014a) “Ensembl Gene: BMPR1A ENSG00000107779” Ensembl. última actualização Agosto de 2014. [Consultado em 05.08.2014].
Disponível em:
http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000107779;r=10:88516407-88692595
- Ensembl, (2014b) “Ensembl Gene: HSPA13 ENSG00000155304” Ensembl. última actualização Agosto de 2014. [Consultado em 05.08.2014].
Disponível em:
http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000155304;r=21:15743436-15755805
- Ensembl, (2014c) “Ensembl Gene: SAMS1 ENSG00000155307” Ensembl. última actualização Agosto de 2014. [Consultado em 05.08.2014].
Disponível em:
http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000155307;r=21:15857549-15955723;t=ENST00000285670
- Forbes, S. a et al., (2011) COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer. *Nucleic acids research*, 39, pp.945–50.
- Francisco, I. et al., (2011) Familial colorectal cancer type X syndrome: two distinct molecular entities? *Familial cancer*, 10(4), pp.623–31.
- Genecards®, (2014a) “Genecards® = BMPR1A”, Genecards®. última actualização Setembro de 2014. [Consultado em 05.08.2014].
Disponível em: <http://www.genecards.org/cgi-bin/carddisp.pl?gene=bmpr1a>
- Genecards®, (2014b) “Genecards® = HSPA13”, Genecards®. última actualização Setembro de 2014. [Consultado em 05.08.2014].
Disponível em: <http://www.genecards.org/cgi-bin/carddisp.pl?gene=hspa13>
- Genecards®, (2014c) “Genecards® = SAMS1” Genecards®. última actualização Setembro de 2014. [Consultado em 05.08.2014].
Disponível em: <http://www.genecards.org/cgi-bin/carddisp.pl?gene=sams1>
- Grizenkova, J. et al., (2012) Overexpression of the Hspa13 (Stch) gene reduces prion disease incubation time in mice. *Proceedings of the National Academy of Sciences of the United States of America*, 109(34), pp.13722–13727.
- Hardwick, J. et al., (2008) Bone morphogenetic protein signalling in colorectal cancer. *Nature reviews. Cancer*, 8, pp.806–812.

- Hardwick, J.C.H. et al., (2004) Bone morphogenetic protein 2 is expressed by, and acts upon, mature epithelial cells in the colon. *Gastroenterology*, 126(1), pp.111–121.
- Hennig, E.E. et al., (2012) Comparative kinome analysis to identify putative colon tumor biomarkers. *Journal of molecular medicine (Berlin, Germany)*, 90(4), pp.447–56.
- Holleben, M. et al., (2011) Immunoinhibitory adapter protein Src homology domain 3 lymphocyte protein 2 (SLy2) regulates actin dynamics and B cell spreading. *The Journal of biological chemistry*, 286(15), pp.13489–501.
- Howe, J.R., (2004) The prevalence of MADH4 and BMPR1A mutations in juvenile polyposis and absence of BMPR2, BMPR1B, and ACVR1 mutations. *Journal of Medical Genetics*, 41(7), pp.484–491.
- Jaspersion, K.W. et al., (2011) Hereditary and Familial Colon Cancer. *Gastroenterology*, 138(6), pp.2044–2058.
- Jemal, A., Bray, F. & Ferlay, J., (2011) Global Cancer Statistics. *American Cancer Society*, 61(2), pp.69–90.
- Jung, D. et al., (2004) Human ileal bile acid transporter gene ASBT (SLC10A2) is transactivated by the glucocorticoid receptor. *Gut*, 53(1), pp.78–84.
- Kaczmarczyk, M. et al., (2013) Association of BMPR1A polymorphism, but not BMP4, with kidney size in full-term newborns. *Pediatric nephrology (Berlin, Germany)*, 28(3), pp.433–8.
- Knösel, T. et al., (2002) Incidence of chromosomal imbalances in advanced colorectal carcinomas and their metastases. *Virchows Archiv*, 440(2), pp.187–194.
- Lapierre, M. et al., (2014) RIP140 increases APC expression and controls intestinal homeostasis and tumorigenesis. *The Journal of Clinic Investigation*, 124(5), pp.1899–1913.
- Leggett, B. & Whitehall, V., (2010) Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology*, 138(6), pp.2088–100.
- Lindor, (2009) Best Practice & Research Clinical Gastroenterology Hereditary colorectal cancer : MYH-associated polyposis and other newly identified disorders. *Best Practice & Research Clinical Gastroenterology*, 23(1), pp.75–87.
- Lindor, N. et al., (2005) Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency. *JAMA*, 293(16), pp.1979–1985.
- Liu, F. et al., (2012) Systematic review and meta-analysis of the relationship between EPHX1 polymorphisms and colorectal cancer risk. *PLOS one*, 7(8), p.e43821.

- Livak, K.J. & Schmittgen, T.D., (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Elsevier Science*, 25(4), pp.402–8.
- Luu, V. et al., (2005) Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction. *BioTechniques*, 38(2), pp.287–93.
- Lynch, H.T. & de la Chapelle, A., (2003) Hereditary colorectal cancer. *The New England journal of medicine*, 348(10), pp.919–32.
- Mayer, M.P. & Bukau, B., (2005) Hsp70 chaperones: cellular functions and molecular mechanism. *Cellular and molecular life sciences : CMLS*, 62(6), pp.670–84.
- McLaren, W. et al., (2010) Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, 26(16), pp.2069–70.
- Nieminen, T.T. et al., (2011) BMPR1A mutations in hereditary nonpolyposis colorectal cancer without mismatch repair deficiency. *Gastroenterology*, 141(1), pp.e23–6.
- Pandurangan, A.K., (2013) Potential Targets for Prevention of Colorectal Cancer : a Focus on PI3K / Akt / mTOR and Wnt Pathways. *Asian pacific Journal of Cancer Prevention*, 14, pp.2201–2205.
- Pereira, C., (2013) “*Estudo do gene NR1P1 e de novos loci de susceptibilidade para o cancro do cólon e recto familiar do tipo X.*” Escola Superior de Saúde Egas Moniz.
- Peters, J. et al., (2011) Molecular ruler of tripeptidylpeptidase II: mechanistic principle of exopeptidase selectivity. *Biochemical and biophysical research communications*, 414(1), pp.209–14.
- Pino, M.S. & Chung, D.C., (2010) The chromosomal instability pathway in colon cancer. *Gastroenterology*, 138(6), pp.2059–72.
- Pires, S., (2011) *Identificação de regiões mínimas de perda de heterozigotia e envolvimento do gene STK24 no cancro do cólon e recto familiar do tipo X.* , Tese de mestrado em Biologia molecular em saúde, Escola Superior de Saúde Egas Moniz, Portugal.
- Póvoa, V., (2011) “*Estudo de genes de susceptibilidade para o cancro do cólon e recto familiar do tipo X.*” , Tese de mestrado em Biologia molecular em saúde, Escola Superior de Saúde Egas Moniz, Portugal.
- Rasmussen, R., (2001) “*Rapid Cycle Real-Time PCR*” *Methods and Applications* 2001st ed., Springer Berlin Heidelberg.

- Roche Applied Science, (2014) “*LightCycler® Real-Time PCR Systems*”, *Application Manual*,
- Saito-diaz, K. et al., (2013) The way Wnt works: Components and mechanism. *Growth Factors*, 31(1), pp.1–31.
- Schnurr, M. et al., (2009) Antigen Delivery and Processing via Tripeptidyl Peptidase II. *The Journal of Immunology*, 182, pp.1253–1259.
- Schwarz, J.M. et al., (2010) MutationTaster evaluates disease-causing potential of sequence alterations. *Nature methods*, 7(8), pp.575–6.
- Sharma, D. & Masison, D.C., (2009) Hsp70 Structure, Function, Regulation and Influence on Yeast Prions. *Protein and peptide letters*, 16(6), pp.571–581.
- Shen, L. et al., (2005) MGMT promoter methylation and field defect in sporadic colorectal cancer. *Journal of the National Cancer Institute*, 97(18), pp.1330–8.
- Sinha, S. & Yang, W., (2008) Cellular signaling for activation of Rho GTPase Cdc42. *Cellular signalling*, 20(11), pp.1927–34.
- Smallridge, R., (2003) “The complexity of protease cascades.” *Nature Reviews Molecular Cell Biology*.
- Smith, P.J. et al., (2006) An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Human molecular genetics*, 15(16), pp.2490–508.
- Tomkinson, B., Ní Laoi, B. & Wellington, K., (2002) The insert within the catalytic domain of tripeptidyl-peptidase II is important for the formation of the active complex. *European journal of biochemistry*, 269(5), pp.1438–43.
- Umar, A. et al., (2004) Testing guidelines for hereditary non-polyposis colorectal cancer. *Nature reviews. Cancer*, 4(2), pp.153–8.
- Vasen, H.F.A. et al., (2008) Guidelines for the clinical management of familial adenomatous polyposis (FAP). *GUT*, pp.704–713.
- Wang, D. et al., (2010) Enhanced adaptive immunity in mice lacking the immunoinhibitory adaptor Hacs1. *The FASEB Journal*, 24(3), pp.947–56.
- Wang, S. & El-Deiry, W.S., (2003) TRAIL and apoptosis induction by TNF-family death receptors. *Oncogene*, 22(53), pp.8628–33.
- Watanabe, T. et al., (2011) Predicting ulcerative colitis-associated colorectal cancer using reverse-transcription polymerase chain reaction analysis. *Clinical colorectal cancer*, 10(2), pp.134–41.
- Wheeler, J. et al., (2013) The Genetics of Keratoconus: A Review. *Reproductive System & Sex Disorders*, pp.1–19.

- Wong, M.H. et al., (1995) Identification of a Mutation in the Ileal Sodium-dependent Bile Acid Transporter Gene That Abolishes Transport Activity. *The Journal of biological chemistry*, 270(45), pp.27228–27234.
- Yamada, H. et al., (2008) Detailed Characterization of a Homozygously Deleted Region Corresponding to a Candidate Tumor Suppressor Locus at 21q11-21 in Human Lung Cancer. *Genes Chromosomes Cancer*, 47(9), pp.810–818.
- Yamagata, N. et al., (2008) Stomach cancer-derived del223V-226L mutation of the STCH gene causes loss of sensitization to TRAIL-mediated apoptosis. *Biochemical and biophysical research communications*, 376(3), pp.499–503.
- Zhao, A., (2012) *Estudo do gene DOCK9 e de novos loci de susceptibilidade no Cancro do Cólon e Recto Familiar do tipo X (FCCTX)*. , Tese de mestrado em Biologia molecular em saúde, Escola Superior de Saúde Egas Moniz, Portugal.
- Zhu, Y.X. et al., (2004) The SH3-SAM adaptor HACCS1 is up-regulated in B cell activation signaling cascades. *The Journal of experimental medicine*, 200(6), pp.737–47.

Anexos

Anexo I. Sequências e temperaturas de *annealing* dos *primers* utilizados para os genes incluídos neste estudo

1) *BMPRIA*

Localização (tamanho)	Sequência (5'→3')	Tm (°C)
Exão 3 (369pb)	<i>Fwd</i> : GGCATTTGTTTCCCTTTTAG	56
	<i>Rev</i> : AAGAACAGAAGCAAGAAATAGTG	62
Exão 4 (574pb)	<i>Fwd</i> : GCTTAACAACCTTTTCTCATT	54
	<i>Rev</i> : TTTTGGTAACCTCTAAGTTTTC	58
Exão 5 (274pb)	<i>Fwd</i> : CCTGTTCACATTCAGACTCA	58
	<i>Rev</i> : TACACCAGGCATGTTTCTC	56
Exão 6 e 7 (547pb)	<i>Fwd</i> : TCACTCACTGAAATAGAAATTG	58
	<i>Rev</i> : AGGAGATATACAACCAGAGAATAC	66
Exão 8 (299pb)	<i>Fwd</i> : TTCTGAGGGAAGGATAATGG	58
	<i>Rev</i> : GAGGACACATTCTAAATCAACAC	64
Exão 9 (327pb)	<i>Fwd</i> : GATGGACTACCCCTTTG	52
	<i>Rev</i> : CCTGTACCTCACTAGCCTTG	62
Exão 10 (524pb)	<i>Fwd</i> : CCTAAGTTTTTCTCAGTATCCAG	64
	<i>Rev</i> : CCCCCAACTATACATTTATG	52
Exão 11 (454pb)	<i>Fwd</i> : AATGCTAAATTCCACAATGC	54
	<i>Rev</i> : TCCTTATCCAATAATATATACACC	58
Exão 12 (280pb)	<i>Fwd</i> : CGTTTTAGTTCCATAGTTTAGC	60
	<i>Rev</i> : TATATTTGGTGGCATTACG	54
Exão 13 (1905pb)	<i>Fwd</i> : ACAGTGATGAAGTGAGTGGAAC	64
	<i>Rev</i> : ATTTTGGTAATGGGTGGAC	54

2) *HSPA13*

Localização (tamanho)	Sequência (5'→3')	Tm (°C)
Exão 1 (249pb)	<i>Fwd:</i> GTTGAATGGAGGCTCTTTGC	60
	<i>Rev:</i> CCGCAAGAGCAACAAGGAC	60
Exão 2 (567pb)	<i>Fwd:</i> TTGCTGTCTGAGAGGAGTGC	62
	<i>Rev:</i> GAAGCCAAAGCATTAATACTAAAAGA	64
Exão 3 (385pb)	<i>Fwd:</i> TTTTGTAAGATAAGGGTAAGTTCTCA	70
	<i>Rev:</i> GAAGGTTTTCAACCCAACACA	60
Exão 4 (355pb)	<i>Fwd:</i> TGTTTCGGTTTGTGACTGCTC	60
	<i>Rev:</i> GATTACAGGCGTGAGCCACT	62
Exão 5 e região 3'UTR (499pb)	<i>Fwd:</i> AAAATTGTAATGATACATGAAACAGG	66
	<i>Rev:</i> GGGTATCAAAGAGTTTCCGTGA	64
Região 3'UTR (496pb)	<i>Fwd:</i> TGACACTGAACTGCCAAAAGA	60
	<i>Rev:</i> TGGGAAGAGATCATCAGACAAG	64

3) *SAMSNI*

Localização (tamanho)	Sequência (5'→3')	Tm (°C)
Exão 2 (984pb)	<i>Fwd:</i> GACACACTGCATCCATACATCTC	66
	<i>Rev:</i> CTATCTGTTTCTTTCCCGACA	64
Exão 3 (500pb)	<i>Fwd:</i> CCAAGGTGCTTACTTAACGCTATG	70
	<i>Rev:</i> GTGACAAGAACCCATTGCTTTCC	68
Exão 4 (483pb)	<i>Fwd:</i> CTGAGTTCGTAACACAAGCAACC	68
	<i>Rev:</i> GTTCTGGGCATAATGACTGAGG	66
Exão 5 (296pb)	<i>Fwd:</i> CATCTCTGTGGAAATCTTCTCTCTG	72
	<i>Rev:</i> CCATTTCTCTTATCAAGTAGCTGGA	70
Exão 6 (344pb)	<i>Fwd:</i> CATCCAGTGTTCAGCATTATC	64
	<i>Rev:</i> CACTGTCTTCCCACTGCTTAT	62
Exão 7 (431pb)	<i>Fwd:</i> CCA CCA GCC TGA ACT TTT GAA	62
	<i>Rev:</i> ACT CCC TTC GGT GTT TCC AT	60

Exão 8 (328pb)	<i>Fwd: CCC ATG TGC ATC TCT ATA CTT TC</i>	66
	<i>Rev: ACT GGG GCT TTG TTT CTA ATG A</i>	62
Exão 9 (554pb)	<i>Fwd: GAT GCG GGA AGT TCA CAT GA</i>	60
	<i>Rev: CTT ATC TTC CTC TCC TAT TTG ACG</i>	68

4) *SLC10A2*

Localização (tamanho)	Sequência (5'→3')	Tm (°C)
Exão 1 (536pb)	<i>Fwd: AAGTTTGATGGGGTAGAAG</i>	54
	<i>Rev: GAAGGATTACTCCAAGGTGA</i>	58
Exão 3 (224pb)	<i>Fwd: AGGCTGCTTGTTCTTCTGTT</i>	58
	<i>Rev: ACCCTATGGTTTTGATTGTC</i>	56
Exão 5 (274pb)	<i>Fwd: CCTAACACTGTGGTAAATGAC</i>	60
	<i>Rev: GATAATATGACGGTGGCTT</i>	54

5) *Primers para amplificação de cDNA*

Localização (tamanho)	Sequência (5'→3')	Tm (°C)
TPP2 transcrito do exão 13 e 14 (Fragmento 5 – 577pb)	<i>Fwd: AATAACCGTGGCATCTACCT</i>	58
	<i>Rev: CTTCACTCAGTGTTTCCTTTCTC</i>	64
TPP2 transcrito do exão 23 e 24 (Fragmento 8 – 528pb)	<i>Fwd: ACCAGCCATTCTTTGTTACT</i>	56
	<i>Rev: TACGAGGGTGGATTTTTGT</i>	54
TEX30 transcrito wt (950pb)	<i>Fwd: GAGATGTCCTTTGCTTCTCA</i>	58
	<i>Rev: ACAACTGTAACAGTGCTTTCAA</i>	60
SAMSN1 transcrito wt (1281pb)	<i>Fwd: GCACACTGCTGACTGTTTTCA</i>	62
	<i>Primer Reverse para análise mutacional do exão 9</i>	
GAPDH transcrito wt (226pb)	<i>Fwd: GAAGGTGAAGGTCGGAGTC</i>	60
	<i>Rev: GAAGATGGTGATGGGATTTC</i>	58

6) *Primers* para estudo de expressão por qPCR

Localização (tamanho)		Sequência (5'→3')	Tm (°C)
<i>HSPA13</i> transcrito <i>wt</i> (109pb)	Exão 3	<i>Fwd</i> : GGAATGCCAGTTGCCAATGCTGTC	74
	Exão 3 e 4	<i>Rev</i> : CCCTCAAAATCTTCAGTCCTGC	66
<i>SAMSNI</i> transcrito <i>wt</i> (162pb)	Exão 2	<i>Fwd</i> : GCACTCACTAGGAGTAGTCAG	64
	Exão 3 e 4	<i>Rev</i> : GTGCCTCAGTTGAATCATCTGG	66
<i>TPP2</i> transcrito <i>wt</i> (93pb)	Exão 13 e 14	<i>Fwd</i> : CCGGAGAACACAGAAACTCTG	66
	Exão 14	<i>Rev</i> : CAAATGGCTGGGACACTGAAC	64
<i>TPP2</i> transcrito alternativo 13a (114pb)	Exão 13	<i>Fwd</i> : GATCATGGCGTTGGCATTGAAC	66
	Exão alternativo 13a	<i>Rev</i> : CCTAATGAGCTGCAGTTCTCC	64
<i>TEX30</i> transcrito <i>wt</i> (64pb)	Exão 3	<i>Fwd</i> : CTTACACATGGAGCATCAGGAG	66
		<i>Rev</i> : CAAGATGGGATGCCAGTGAC	62
<i>SDHA</i> transcrito <i>wt</i> (90pb)	Exão 13 e 14	<i>Fwd</i> : TACAAGGTGCGGATTGATGA	58
	Exão 14	<i>Rev</i> : GGTGTGCTTCCTCCAGTGCT	64
<i>RPS13</i> transcrito <i>wt</i> (95pb)	Exão 2	<i>Fwd</i> : CAGTCGGCTTTACCCTATCG	62
	Exão 3	<i>Rev</i> : CCCTTCTTGGCCAGTTTGTA	60

7) Primers para estudo de copy number por qPCR

Localização (tamanho)		Sequência (5'→3')	Tm (°C)
HSPA13 (116pb)	Exão 3	<i>Primer forward</i> para análise de expressão	
	Intrão 3	<i>Rev:</i> GCATCTCAGTCAAGGGTACTG	64
SAMSNI (132pb)	Exão 5	<i>Fwd:</i> GGATAGTCTCTACAGTGGACAG	66
	Intrão 5	<i>Rev:</i> CAAGTAGCTGGAACCTTGTGG	64
TPP2 (65pb)	Intrão 6 e exão 7	<i>Fwd:</i> AGGAGCTCATGGGACACATGT	64
	Exão 7	<i>Rev:</i> TCCGTTTCAGGTTCTTCTGGAA	62
TPP2 (69pb)	Intrão 21	<i>Fwd:</i> AAATCACTGCTTTTGTCAATTCTCAA	66
	Intrão 21	<i>Rev:</i> AACGACAGAATACTCCCACACTTTT	70
DOCK9 (65pb)	Intrão 2	<i>Fwd:</i> TGGTGACTGGCTTCCAAGAAG	64
	Intrão 2	<i>Rev:</i> GCCCAAATCTAGCCTTTAGAAAAA	66
CHR12 (100pb)		<i>Fwd:</i> GCTTGCCCTGTCCAGTTAAT	60
(Região cromossômica 6585217-6585317)		<i>Rev:</i> TAGCTCAGCTGCACCCTTTA	60

Anexo II. Condições de PCR utilizados na amplificação dos exões dos genes incluídos neste estudo

1) Condições de PCR para amplificação dos exões do gene *BMPRIA*

Localização	Temperatura (°C)	MgCl ₂ (μl)	<i>primer</i> (μl)	Ciclos	Reagentes
Exão 3	61	0,75	0,15	35	Bioline
Exão 4	52,5	0,75	0,15	35	Bioline
Exão 5	58	0,5	0,15	35	Bioline
Exões 6/7	62	1	0,15	35	Bioline
Exão 8	58	1	0,15	35	Bioline
Exão 9	63	1,75	0,15	35	Bioline
Exão 10	61,5	1	0,15	35	Bioline
Exão 11	60	1,25	0,15	35	Bioline
Exão 12	56	1	0,15	35	Bioline
Exão 13 e região 3'UTR	62	1,25	0,2	40	AmpliTaq Gold (Applied Biosystems)

2) Condições de PCR para amplificação dos exões do gene *HSPA13*

Localização	Temperatura (°C)	MgCl ₂ (μl)	DMSO	<i>primer</i> (μl)	Ciclos	Reagentes
Exão 1	65	0,75	-	0,3	35	Bioline
Exão 2	67	1,25	-	0,3	35	Bioline
Exão 3	64	1,75	-	0,3	35	AmpliTaq Gold (Applied Biosystems)
Exão 4	61,5	1,5	4%	0,3	35	Bioline
Exão 5 e região 3'UTR	60	1,5	4%	0,3	35	AmpliTaq Gold (Applied Biosystems)
Região 3'UTR	64,5	1	-	0,3	35	Bioline

3) Condições de PCR para amplificação dos exões do gene *SAMSNI*

Localização	Temperatura (°C)	MgCl ₂ (μl)	DMSO	<i>primer</i> (μl)	Ciclos	Reagentes
Exão 2	71	1,25	—	0,3	35	Bioline
Exão 3	71	0,75	—	0,3	35	Bioline
Exão 4	69	0,75	—	0,3	35	Bioline
Exões 5	70	1,75	—	0,3	35	Bioline
Exão 6	70	1	—	0,3	35	Bioline
Exão 7	66	0,75	—	0,3	35	Bioline
Exão 8	66	1,5	4%	0,3	35	Bioline
Exão 9	66	1,5	4%	0,3	35	Bioline

4) Condições de PCR para amplificação dos exões do gene *SLC10A2*

Localização	Temperatura (°C)	MgCl ₂ (μl)	<i>primer</i> (μl)	Ciclos	Reagentes
Exão 1	64	1	0,3	35	Bioline
Exão 3	59,5	1	0,3	35	Bioline
Exão 5	63	1,25	0,3	35	Bioline

5) Condições de PCR para amplificação de cDNA

cDNA	Temperatura (°C)	MgCl ₂ (μl)	<i>primer</i> (μl)	Ciclos	Reagentes
<i>TEX30</i> transcrito <i>wt</i>	61,5	0,5	0,3	40	Platinum
<i>TPP2</i> transcrito do exão 13 e 14	55	1,25	0,9	35	Bioline
<i>TPP2</i> transcrito do exão 23 e 24	53,5	2,5	0,3	35	Amplitaq Gold (Applied Biosystems)
<i>SAMSNI</i> transcrito <i>wt</i>	64	1,25	0,3	35	Bioline
<i>GAPDH</i> transcrito <i>wt</i>	49	1,5	0,3	35	Invitrogen

Anexo III. Programas de PCR utilizados na amplificação dos exões dos genes incluídos neste estudo

1) Programa utilizado na amplificação por PCR dos genes presentes neste estudo utilizando os reagentes do *kit Bioline*

Etapa	Temperatura (°C)	Tempo	Número de ciclos
Desnaturação inicial	95	3 min	1
Desnaturação	94	50 seg	35 a 40
<i>Annealing</i>	Variável	30 seg	
Extensão	72	50 seg	
Extensão final	72	7 min	1
Pausa	15	-	-

2) Programa utilizado na amplificação por PCR dos genes presentes neste estudo utilizando os reagentes do *kit AmpliTaq Gold* (Applied Biosystems)

Etapa	Temperatura (°C)	Tempo	Número de ciclos
Desnaturação inicial	95	7 min	1
Desnaturação	95	25 seg	40
<i>Annealing</i>	Variável	25 seg	
Extensão	70	1 min	
Extensão final	70	10 min	1
Pausa	15	-	-

3) Programa utilizado na amplificação por PCR dos genes presentes neste estudo utilizando os reagentes do *kit Platinum PCRx* (Invitrogen)

Etapa	Temperatura (°C)	Tempo	Número de ciclos
Desnaturação inicial	94	30 seg	1
Desnaturação	94	30 seg	40
<i>Annealing</i>	Variável	30 seg	
Extensão	72	1 min	
Pausa	15	-	-

4) Programa utilizado na amplificação por PCR de cDNA (RT-PCR)

Etapa	Temperatura (°C)	Tempo	Número de ciclos
Desnaturação e ligação dos hexâmeros	70	10 min	1
Pausa	10	∞	
Síntese de cDNA	42	60 min	
Extensão	70	15 min	
Pausa	4	∞	-

Anexo IV. Condições e programa de qPCR utilizados na amplificação dos genes incluídos neste estudo

1) Programa utilizado na amplificação por qPCR

Etapa	Número de ciclos	Modo de análise
Pre-incubação	1	Nenhum
Amplificação	45	Quantificação
Curvas de <i>melting</i>	1	Curva de <i>melting</i>
Arrefecimento	1	Nenhum

Etapa	Temperatura (°C)	Modo de aquisição	Tempo	Ramp rate (°C/s)	Aquisições (por °C)
Pre-incubação	95	Nenhum	10 min	4,4	—
Amplificação	95	Nenhum	10 seg	4,4	—
	65	Nenhum	15 seg	2,2	—
	72	Simples	35 seg	4,4	—
Curvas de <i>melting</i>	95	Nenhum	5 seg	4,4	—
	65	Nenhum	1 min	2,2	—
	97	Contínuo	—	—	10
Arrefecimento	40	Nenhum	10 seg	1,5	—

2) Condições de qPCR para estudo de expressão

Localização	[cada <i>primer</i>] (μ mol)	Reagentes
<i>SAMSNI</i> transcrito <i>wt</i>	4	SYBR Green I Master (Roche)
<i>HSPA13</i> transcrito <i>wt</i>	6	SYBR Green I Master (Roche)
<i>TEX30</i> transcrito <i>wt</i>	5	SYBR Green I Master (Roche)
<i>TPP2</i> transcrito <i>wt</i>	5	SYBR Green I Master (Roche)
<i>TPP2</i> transcrito alternative 13a	5	SYBR Green I Master (Roche)
<i>SDHA</i> transcrito <i>wt</i>	5	SYBR Green I Master (Roche)
<i>RPS13</i> transcrito <i>wt</i>	5	SYBR Green I Master (Roche)

3) Condições de qPCR para estudo de *copy number* em sangue periférico

Localização	[cada <i>primer</i>] (μ mol)	Reagentes
<i>SAMSNI</i>	2	SYBR Green I Master (Roche)
<i>HSPA13</i>	4	SYBR Green I Master (Roche)
<i>TPP2</i> intrão 6	5	SYBR Green I Master (Roche)
<i>TPP2</i> intrão 21	3	SYBR Green I Master (Roche)
<i>CHR12</i>	10	SYBR Green I Master (Roche)

4) Condições de qPCR para estudo de *copy number* em tecido incluído em parafina

Localização	[cada <i>primer</i>] (μmol)	Reagentes
<i>TPP2</i> intrão 6	5	SYBR Green I Master (Roche)
<i>TPP2</i> intrão 21	8	SYBR Green I Master (Roche)
<i>DOCK9</i>	8	SYBR Green I Master (Roche)
<i>CHR12</i>	5	SYBR Green I Master (Roche)

Anexo V. Preparação dos reagentes utilizados na análise electroforética dos produtos de PCR

1) Preparação do tampão de electroforese

TBE 1x: diluído a partir de TBE 10x (0,89M Tris Borato pH8,3 + 20mM Na₂ EDTA – *National Diagnostics*) para um volume final de 2000ml, em água bidestilada.

2) Preparação do gel de agarose a 2% (p/v)

1. Pesar 5gr de agarose (*Seaken® LE Agarose* – Lonza), para um *erlenmeyer* de 500ml;
2. Adicionar 250ml de tampão TBE 1x;
3. Dissolver a solução no micro-ondas;
4. Adicionar 12,5µl de brometo de etídio (10mg/ml, MP biomedical) e agitar para homogeneização;
5. Colocar a solução num molde com 4 pentes e deixar arrefecer até à solidificação do gel.

Durante este estudo foi necessário utilizar outras concentrações ou volumes de gel de agarose apresentando-se no quadro seguinte:

Concentração de agarose	Volume de tampão (ml)	Quantidade de agarose (gr)	Volume de brometo de etídio (µl)
0,8%	150	1,2	7,5
	250	2	12,5
1,2%	150	2,25	7,5
	250	3,75	12,5
1,5%	150	1,8	7,5
	250	3	12,5

3) Preparação do *Orange G* (Sigma) 5x e 1x

Orange G 5x: preparado a partir de 12ml de Ficoll (Sigma), 125gr de *Orange G* (Sigma) e 50ml de água bidestilada. Dissolver com agitação.

Orange G 1x: diluído a partir de *Orange G* 5x (Sigma) para um volume final de 1ml em água bidestilada.

4) Preparação do marcador de peso molecular *GeneRuler 50bp DNA Ladder* (Thermo Scientific)

Adicionar 50µl de *GeneRuler 50bp DNA Ladder* (Thermo Scientific) a 250µl de *Orange G 5x* e perfazer o volume para 1ml com água bidestilada.

5) Preparação do marcador de peso molecular 1kb DNA Ladder (*Promega*)

Adicionar 20µl de 1kb DNA Ladder (*Promega*) a 20µl de *Orange G 5x* e perfazer o volume para 80µl com água bidestilada.

Anexo VI. Protocolo de extracção e purificação de DNA a partir de banda em gel de agarose – *QIAquick Gel Extraction Kit* (Qiagen)

1. Excisar a banda de interesse do gel de agarose, com uma lâmina de bisturi limpa e coloca-la dentro de um tubo de 1,5ml vazio;
2. Adicionar 300µl de tampão QG ao tubo de 1,5ml que contém a banda excisada;
3. Incubar a 50°C, durante 10min, misturando a cada 2-3minutos, para ajudar a dissolver a agarose;
4. Adicionar 100µl de isopropanol e agitar no vortex;
5. Colocar uma coluna num tubo colector de 2ml;
6. Adicionar todo o volume, contido no tubo de 1,5ml à coluna e centrifugar a 13000rpm, durante 1 min;
7. Descartar o volume contido no tubo colector e recolocar a coluna no mesmo;
8. Adicionar 500µl de tampão QG à coluna e centrifugar a 13000rpm, durante 1 min;
9. Descartar o volume contido no tubo colector e recolocar a coluna no mesmo;
10. Adicionar 750µl de tampão PE à coluna, incubar 2-5min à temperatura ambiente e centrifugar a 13000rpm, durante 1min;
11. Descartar o volume do tubo colector e centrifugar novamente durante 1 min;
12. Colocar a coluna num novo tubo de 1,5ml;
13. Adicionar 30µl de tampão EB ao centro da membrana da coluna e incubar durante 1 min à temperatura ambiente;
14. Centrifugar durante 1 min a 13000 rpm, para eluir o produto purificado;
15. Armazenar o produto purificado a -20°C.

Anexo VII. Protocolo de precipitação e purificação do DNA
Etanol/EDTA/Acetato de sódio – *BigDye® Terminator v1.1 Cycle Sequencing Kit* (Applied Biosystems) após a reacção de sequenciação

1. Num tubo de 1,5ml, devidamente identificado adicionar 2µl de EDTA (125mM), 2µl de acetato de sódio (3M) e 50µl de etanol absoluto;
2. Adicionar o produto da reacção de sequenciação e agitar em vortex;
3. Incubar durante 15min à temperatura ambiente;
4. Centrifugar a 14000rpm durante 30min a 4°C;
5. Retirar completamente o sobrenadante com uma micropipeta;
6. Adicionar 100µl de etanol a 70%(v/v) e agitar ligeiramente em vortex;
7. Centrifugar a 14000rpm durante 15min a 4°C;
8. Retirar completamente o sobrenadante;
9. Secar o pellet a 37°C durante 10min;
10. Guardar o pellet a 4°C.

Anexo VII. Resultados

1) Análise mutacional do gene *HSPA13* em indivíduos de famílias FCCTX

Família	DNA	Exão 1 e região flanqueadora	Exão 2 e região flanqueadora	Exão 3 e região flanqueadora	Exão 4 e região flanqueadora	Exão 5, região flanqueadora e região 3'UTR
7	L5	N	N	N	Het	N
24	L48	N	N	N	Het	N
37	L108	N	N	N	Homo	N
55	L295	N	N	N	Homo	N
	L926	N	N	N	N	N
	L950	N	N	N	N	N
56	L801	N	N	N	Homo	N
63	L451	N	N	N	Het	N
69	E436	N	N	N	Het	N
71	L467	N	N	N	Homo	N
84	L484	N	N	N	Homo	N
86	L499	N	N	N	Homo	N
106	L701	N	N	N	Het	Het
148	L984	N	N	N	Het	N
173	L1038	N	N	N	Homo	N
236	L1200	N	N	N	N	N
P2209	L1121	N	N	N	N	N

Exão 4 e região flanqueadora: c.581-46 G>T; Exão 5 e região 3'UTR: c.749-56 G>A.

2) Análise mutacional do gene *SAMSN1* em indivíduos de famílias FCCTX

Família	DNA	Exão 2 e região flanqueadora	Exão 3 e região flanqueadora	Exão 4 e região flanqueadora	Exão 5 e região flanqueadora	Exão 6 e região flanqueadora	Exão 7 e região flanqueadora	Exão 8 e região flanqueadora	Exão 9, região flanqueadora e região 3'UTR
7	L5	Homo 2; Homo3	N	N	N	N	N	N	N
24	L48	Homo 3	Het 1; Het 2	N	N	N	N	N	N
37	L108	Homo 3	N	N	N	N	N	N	Het
55	L295	Homo 1; Homo 3	Het 1; Het 2	N	N	N	N	N	N
	L926	Homo 1; Homo 3	Het 1; Het 2	N	N	N	N	N	N
	L950	Het 1; Het 2; Homo3	Het 1; Het 2	N	N	N	N	N	Het
56	L801	Het 1; Het 2; Homo3	N	N	N	N	N	N	N
63	L451	Het 1; Het 2; Homo 3	N	N	N	N	N	N	Het
69	E436	Homo 2; Homo3	Het 1	N	N	N	N	N	N
71	L467	Het 1; Het 2; Homo 3	Het 1	N	N	N	N	N	N
84	L484	Het 3	N	N	N	N	N	N	N
86	L499	Homo 1; Homo 3	Het 1; Het 2	N	N	N	N	N	N
106	L701	Homo 2; Homo3	N	N	N	N	N	N	N
148	L984	Het 1; Het 3	Het 1	N	N	N	N	N	N
173	L1038	Het 1; Het 2; Homo 3	Homo 1; Het 2	N	N	N	N	N	N
236	L1200	Homo 2; Homo3	N	N	N	N	N	N	N
P2209	L1121	Het 3	Het 1; Het 2	N	N	N	N	N	N

Exão 2 e região flanqueadora: 1 – c.71 C>T, 2 – c.190 C>T, 3 – c.58 A>G; Exão 3 e região flanqueadora: 1 – c.262-52 A>G, 2 – c.262-25 T>A; Exão 9, região flanqueadora e região 3'UTR: c.920-124 C>T.

3) Análise mutacional do gene *BMPRIA* em indivíduos de famílias FCCTX

Família	DNA	Exão 3 e região flanqueadora	Exão 4 e região flanqueadora	Exão 5 e região flanqueadora	Exão 6 e 7 e regiões flanqueadoras	Exão 8 e região flanqueadora	Exão 9 e região flanqueadora	Exão 10 e região flanqueadora	Exão 11 e região flanqueadora	Exão 12 e região flanqueadora	Exão 13, região flanqueadora e região 3'UTR
7	L5	N	N	N	N	N	N	N	N	N	Het 2; Het 3
24	L48	N	N	N	N	N	N	N	N	N	Het 2; Homo 3
37	L108	Het	N	N	Het	N	N	N	N	Het	N
55	L295	Het	N	N	N	N	N	N	N	Het	Het 4
	L926	Het	N	N	N	N	N	N	N	Het	Het 4
	L950	N	N	N	N	N	N	N	N	N	Het 2
56	L801	N	N	N	N	N	N	N	N	N	N
63	L451	N	N	N	N	N	N	N	N	N	N
69	E436	N	N	N	N	N	N	N	N	N	Het 2; Het 3
71	L467	Het	N	N	N	N	N	N	N	Het	Het 1; Het 5
84	L484	N	N	N	N	N	N	N	N	N	N
86	L499	Het	N	N	N	N	N	N	N	Het	Het 4
106	L701	Het	N	N	N	N	N	N	N	Het	N
148	L984	Het	N	N	N	N	N	N	N	Het	Het 2; Het 5
173	L1038	N	N	N	N	N	N	N	N	N	N
236	L1200	N	N	N	N	N	N	Het	N	N	Het 2
P2209	L1121	Het	N	N	N	N	N	N	N	Het	Het 2; Het 5

Exão 3 e região flanqueadora: c.4 C>A; Exão 6 e 7 e regiões flanqueadoras: c.431-26 T>A; Exão 10 e região flanqueadora: c.1140 C>T; Exão 12 e região flanqueadora: c.448-11 T>C; Exão 13, região flanqueadora e região 3'UTR: 1 – c.1599+847 G>A, 2 – c.1599+886 C>T, 3 – c.1599+1497 T>C, 4 – c.1599+764 T>C, 5 – c.1577_delTTTA.